



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 151833

TO: Minh-Tam Davis
Location: REM-3A24&3C18
Art Unit: 1642
Friday, April 29, 2005

Case Serial Number: 09/674237

From: Toby Port
Location: Biotech-Chem Library
REM1-A59
Phone: 272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Davis,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

proty 02/99
04/98

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

151838

From: Chan, Christina
Sent: Wednesday, April 27, 2005 12:34 PM
To: Davis, Minh-Tam; STIC-Biotech/ChemLib
Subject: RE: Rush search request for 09/674237

Please ~~rush~~. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

CHT

-----Original Message-----

From: Davis, Minh-Tam
Sent: Wednesday, April 27, 2005 12:31 PM
To: Chan, Christina
Subject: Rush search request for 09/674237

Please search in commercial database, issued patent files, and PGPUB:

- 1) A nucleotide sequence encoding SEQ ID NO:3, with size limitation of the sequences in the database 100 nucleotides or less.
- 2) SEQ ID NO:1, with size limitation of the sequences in the database 100 nucleotides or less.
- 3) SEQ ID NO:2, with size limitation of the sequences in the database 100 nucleotides or less.

Thank you.

MINH TAM DAVIS
ART UNIT 1642, ROOM 3A24, MB 3C18
272-0830

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIS: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 06:15:47 ; Search time 21159.3 Seconds
(without alignments)
11642.487 Million cell updates/sec

Title: US-09-674-237B-1
Perfect score: 5084
Sequence: 1 cgcacgagaggagtgaggtgaggtgaggtgaggtgaggtt 5084

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 2238514

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	1.3	65	6	CQ557589 Sequence
2	54	1.1	65	6	CQ531908 Sequence
3	48.8	1.0	60	6	CQ541131 Sequence
C 4	41.8	0.8	98	6	BD270475 Synthetic
C 5	41.8	0.8	98	6	AX039308 Sequence
C 6	41.8	0.8	98	6	AX039517 Sequence
7	41.8	0.8	100	6	BD270476 Synthetic
8	41.8	0.8	100	6	AX039309 Sequence
9	41.8	0.8	100	6	AX039518 Sequence
C 10	40.4	0.8	92	6	BD270477 Synthetic
C 11	40.4	0.8	92	6	AX039310 Sequence
C 12	40.4	0.8	92	6	AX039519 Sequence
13	40.4	0.8	94	6	BD270478 Synthetic
14	40.4	0.8	94	6	AX039311 Sequence
15	40.4	0.8	94	6	AX039520 Sequence
16	38	0.7	78	6	AR159552 Sequence
17	37.8	0.7	82	5	HCH388053 Hyla chry
C 18	37.8	0.7	92	8	AF143273 Vitis vin
19	37	0.7	88	6	BD270480 Synthetic

20	37	0.7	88	6	AX039313	Sequence
21	37	0.7	88	6	AX039522	Sequence
22	36.6	0.7	87	6	AX384806	Sequence
23	36.4	0.7	78	6	AR159551	Sequence
C 24	35.8	0.7	86	6	BD270479	Synthetic
C 25	35.8	0.7	86	6	AX039312	Sequence
C 26	35.8	0.7	86	6	AX039521	Sequence
C 27	35.2	0.7	92	9	AY225366	Pan trogl
28	34.4	0.7	99	10	RRGR99	X69669 R. rattus
29	34	0.7	75	6	AR159550	Sequence
C 30	33.4	0.7	93	6	CQ080578	Sequence
C 31	33.4	0.7	93	6	CQ114850	Sequence
C 32	33.4	0.7	93	6	CQ153723	Sequence
C 33	33.4	0.7	93	6	CQ186430	Sequence
C 34	33.4	0.7	93	6	CQ236983	Sequence
C 35	33.4	0.7	93	6	CQ274588	Sequence
C 36	33.4	0.7	93	6	CQ311642	Sequence
C 37	33.4	0.7	93	6	CQ348908	Sequence
38	33.2	0.7	77	9	AY225367	Go rilla g
39	33	0.6	66	6	A62702	Sequence 3
40	33	0.6	69	6	AR159553	Sequence
C 41	32.4	0.6	87	6	CQ877372	Sequence
C 42	31.4	0.6	77	6	BD270481	Synthetic
C 43	31.4	0.6	77	6	AX039314	Sequence
C 44	31.4	0.6	77	6	AX039523	Sequence
45	31.4	0.6	79	6	BD270482	Synthetic

ALIGNMENTS

RESULT 1	CQ557589	65 bp	DNA	linear	PAT 30-JAN-2004
LOCUS	Sequence. 27224 from Patent WO0210449.				
DEFINITION	CQ557589				
ACCESSION	CQ557589.1	GI:41524016			
VERSION					
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
REFERENCE	1				
AUTHORS	Shoshan,A., Wasserman,A., Mintz,E., Mintz,L. and Faigler,S.				
TITLE	Oligonucleotide library for detecting rna transcripts and splice variants that populate a transcriptome				
JOURNAL	Patent: WO 0210449-A 27224 07-FEB-2002;				
FEATURES	Compugen Inc. (US)				
source	Location/Qualifiers				
	1..65				
	/organism="Mus musculus"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:10090"				

ORIGIN

Query Match	1.3%;	Score 65;	DB 6;	Length 65;
Best Local Similarity	100.0%;	Pred. No. 0.00092;		
Matches 65;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	4058	GCAGTTTACCTCATTTGACCTTAGTTGCATGTCGAAATGCTGAGTCACGTCGCA	4117	
Db	1	GCAGTTTACCTCATTTGACCTTAGTTGCATGTCGAAATGCTGAGTCACGTCGCA	60	
QY	4118	GAGGC 4122		
Db	61	GAGGC 65		
RESULT 2	CQ531908	65 bp	DNA	linear
LOCUS	Sequence 1543 from Patent WO0210449.			
DEFINITION	CQ531908			
ACCESSION	CQ531908			


```
QY 1444 GAGCGCAGGAGCAGGAGCCAGCGCAGCTGAGCTGGAGAGCAGCTGGAGAGCAG 1503
      |||||
Db 88 GAGCAGCAGGAGCAGCAGGAGCAGTAGCAGGAGCAGCAGGAGCAGCAGGAGCAGGAG 29
      |||||
QY 1504 CGGAGCTGGAGCGCGCAGCGAGGAGG 1528
      |||||
Db 28 CAGCAGGAGCAGCAGCAGCAGCAGCAG 4

RESULT 6
AX039517/c
LOCUS AX039517 98 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 9 from Patent WO0063374.
ACCESSION AX039517
VERSION AX039517.1 GI:11229544
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Lawson,A.D. and Finney,H.M.
TITLE Synthetic transmembrane components
JOURNAL Patent: WO 0063374-A 9 26-OCT-2000;
CELLTECH THERAPEUTICS LIMITED (GB)
FEATURES
    source
    Location/Qualifiers
        1..98
        /organism="synthetic construct"
        /mol_type="unassigned DNA"
        /db_xref="taxon:32630"
        /note="B6463"
ORIGIN
Query Match 0.8%; Score 41.8; DB 6; Length 98;
Best Local Similarity 68.2%; Pred. No. 2.1e+02;
Matches 58; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 1444 GAGCGCAGGAGCAGGAGCCAGCGCAGCTGAGCTGGAGAGCAGCTGGAGAGCAG 1503
      |||||
Db 88 GAGCAGCAGGAGCAGCAGGAGCAGTAGCAGGAGCAGCAGGAGCAGCAGGAGCAGGAG 29
      |||||
QY 1504 CGGAGCTGGAGCGCGCAGCGAGGAGG 1528
      |||||
Db 28 CAGCAGGAGCAGCAGCAGCAGCAGCAG 4

RESULT 7
BD270476
LOCUS BD270476 100 bp DNA linear PAT 17-JUL-2003
DEFINITION Synthetic transmembrane components.
ACCESSION BD270476
VERSION BD270476.1 GI:33080244
KEYWORDS JP 2002541845-A/10.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 100)
AUTHORS Lawson,A.D.G. and Finney,H.M.
TITLE Synthetic transmembrane components
JOURNAL Patent: JP 2002541845-A 10 10-DEC-2002;
CELLTECH THERAPEUTICS LTD
OS Artificial Sequence
PN JP 2002541845-A/10
PD 10-DEC-2002
PF 17-APR-2000 JP 2000612453
PR 16-APR-1999 GB 9908816.3,16-APR-1999 GB 9908818.9 PI
ALASTAIR DAVID GRIFFITHS LAWSON,HELENE MARGARET FINNEY PC
C12N15/09,A61K38/00,A61K48/00,A61P1/00,A61P3/10,A61P7/06, PC
A61P11/06,
PC A61P17/00,A61P17/06,A61P19/02,A61P25/00,A61P29/00,A61P31/18,
PC A61P35/00,
PC A61P37/06,A61P37/08,C07K14/705,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC C12N15/00,C12N5/00,A61K37/02
```

```
CC B6464
FH Key Location/Qualifiers
FT source 1..100
      /organism="Artificial Sequence".
FEATURES
    source
    Location/Qualifiers
        1..100
        /organism="synthetic construct"
        /mol_type="genomic DNA"
        /db_xref="taxon:32630"
ORIGIN
Query Match 0.8%; Score 41.8; DB 6; Length 100;
Best Local Similarity 68.2%; Pred. No. 2.1e+02;
Matches 58; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 1444 GAGCGCAGGAGCAGGAGCCAGCGCAGCTGAGCTGGAGAGCAGCTGGAGAGCAG 1503
      |||||
Db 15 GAGCAGCAGGAGCAGCAGGAGCAGTAGCAGGAGCAGCAGGAGCAGCAGGAGCAGGAG 74
      |||||
QY 1504 CGGAGCTGGAGCGCGCAGCGAGGAGG 1528
      |||||
Db 75 CAGCAGGAGCAGCAGCAGCAGCAGCAG 99

RESULT 8
AX039309
LOCUS AX039309 100 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 10 from Patent WO0063373.
ACCESSION AX039309
VERSION AX039309.1 GI:11229414
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Finney,H.M. and Lawson,A.D.
TITLE A method of altering the properties of a membrane-associated prote
in by substitution of the transmembrane domain
JOURNAL Patent: WO 0063373-A 10 26-OCT-2000;
CELLTECH THERAPEUTICS LIMITED (GB)
FEATURES
    source
    Location/Qualifiers
        1..100
        /organism="synthetic construct"
        /mol_type="unassigned DNA"
        /db_xref="taxon:32630"
        /note="B6464"
ORIGIN
Query Match 0.8%; Score 41.8; DB 6; Length 100;
Best Local Similarity 68.2%; Pred. No. 2.1e+02;
Matches 58; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 1444 GAGCGCAGGAGCAGGAGCCAGCGCAGCTGAGCTGGAGAGCAGCTGGAGAGCAG 1503
      |||||
Db 15 GAGCAGCAGGAGCAGCAGGAGCAGTAGCAGGAGCAGCAGGAGCAGCAGGAGCAGGAG 74
      |||||
QY 1504 CGGAGCTGGAGCGCGCAGCGAGGAGG 1528
      |||||
Db 75 CAGCAGGAGCAGCAGCAGCAGCAGCAG 99

RESULT 9
AX039518
LOCUS AX039518 100 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 10 from Patent WO0063374.
ACCESSION AX039518
VERSION AX039518.1 GI:11229545
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Lawson,A.D. and Finney,H.M.
```



```
Db 26 GCAGGAGCAGCAGC 13

RESULT 13
BD270478
LOCUS BD270478 94 bp DNA linear PAT 17-JUL-2003
DEFINITION Synthetic transmembrane components.
ACCESSION BD270478
VERSION BD270478.1 GI:33080246
KEYWORDS JP 2002541845-A/12.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
1 (bases 1 to 94)
REFERENCE Lawson,A.D.G. and Finney,H.M.
AUTHORS Synthetic transmembrane components
TITLE Patent: JP 2002541845-A 12 10-DEC-2002;
JOURNAL CELLTech THERAPEUTICS LTD
COMMENT OS Artificial Sequence
PN JP 2002541845-A/12
PD 10-DEC-2002
PF 17-APR-2000 JP 2000612453
PR 16-APR-1999 GB 9908816.3,16-APR-1999 GB 9908818.9 PI
ALASTAIR DAVID GRIFFITHS LAWSON,HELENE MARGARET FINNEY PC
C12N15/09,A61K38/00,A61K48/00,A61P1/00,A61P3/10,A61P7/06, PC
A61P11/06,
PC A61P17/00,A61P17/06,A61P19/02,A61P25/00,A61P29/00,A61P31/18,
PC A61P35/00,
PC A61P37/06,A61P37/08,C07K14/705,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC C12N15/00,C12N5/00,A61K37/02
CC B6466
FH Key Location/Qualifiers
FT source 1..94
FEATURES
source Location/Qualifiers
1..94
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 0.8%; Score 40.4; DB 6; Length 94;
Best Local Similarity 71.6%; Pred. No. 4.3e+02;
Matches 53; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1437 AGAGCGGAGCGCCAGCAGCAGCAGCGCCAGCGCAGCTGGAGCTGGAGAGCAGCTGGA 1496
Db 11 AGAACAGGAGCAGCAGCAGCAGTAGCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCA 70
QY 1497 GAAGCAGCGGAGC 1510
Db 71 GCAGGAGCAGCAGC 84

RESULT 14
AX039311
LOCUS AX039311 94 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 12 from Patent WO0063373.
ACCESSION AX039311
VERSION AX039311.1 GI:11229416
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
1
REFERENCE Finney,H.M. and Lawson,A.D.
AUTHORS A method of altering the properties of a membrane-associated prote
TITLE in by substitution of the transmembrane domain
JOURNAL Patent: WO 0063373-A 12 26-OCT-2000;
CELLTECH THERAPEUTICS LIMITED (GB)
FEATURES
source Location/Qualifiers
1..94
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="B6466"

ORIGIN
Query Match 0.8%; Score 40.4; DB 6; Length 94;
Best Local Similarity 71.6%; Pred. No. 4.3e+02;
Matches 53; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1437 AGAGCGGAGCGCCAGCAGCAGCAGCGCCAGCGCAGCTGGAGCTGGAGAGCAGCTGGA 1496
Db 11 AGAACAGGAGCAGCAGCAGCAGTAGCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCA 70
QY 1497 GAAGCAGCGGAGC 1510
Db 71 GCAGGAGCAGCAGC 84

Search completed: April 28, 2005, 18:48:47
Job time : 21164.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 05:12:48 ; Search time 2462.18 Seconds
(without alignments)
12223.299 Million cell updates/sec

Title: US-09-674-237B-1
Perfect score: 5084
Sequence: 1 cgcacgagagagtggtgag.....gagaattcatatcaagctt 5084

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 4530610

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	1.6	80	ADP49483	Adp49483 Oligonucl
2	65	1.3	65	ABN54476	Abn54476 Mouse spl
3	54	1.1	65	ABN28795	Abn28795 Rat splic
4	48.8	1.0	60	ABN38018	Abn38018 Human spl
5	42.8	0.8	81	ADC16446	Adc16446 Short int
6	42.8	0.8	81	ADC16445	Adc16445 Short int
7	41.8	0.8	98	AAC5404	Aac5404 Oligonucl
8	41.8	0.8	100	AAC5405	Aac5405 Oligonucl
9	41	0.8	75	ABK99299	Abk99299 Trinucleo
10	40.4	0.8	92	AAC5406	Aac5406 Oligonucl
11	40.4	0.8	94	AAC5407	Aac5407 Oligonucl
12	39.2	0.8	97	AAC5408	Aac5408 Human col
13	39	0.8	89	AAV58382	Av58382 Clone #8
14	38	0.7	78	AAV17232	Av17232 SCA2 gene
15	38	0.7	78	AAV17233	Av17233 SCA2 gene
16	37.6	0.7	69	ADS31430	Ad31430 Human gen
17	37	0.7	88	AAC5409	Aac5409 Oligonucl
18	36.6	0.7	87	AAD30431	Ad30431 29 mer ol
19	36.4	0.7	78	AAV17231	Av17231 SCA2 gene
20	35.8	0.7	86	AAC5408	Aac5408 Oligonucl

21	34	0.7	75	2	AAV17230	Aav17230 SCA2 gene
22	33.4	0.7	93	4	AAI26445	Aai26445 Probe #16
23	33.4	0.7	93	4	ABA74544	Aba74544 Human foe
24	33.4	0.7	93	4	AAI55023	Aai55023 Probe #23
25	33.4	0.7	93	4	ABA39360	Aba39360 Probe #17
26	33.4	0.7	93	4	AAK49188	Aak49188 Human bon
27	33.4	0.7	93	4	AAK23011	Aak23011 Human bra
28	33.4	0.7	93	4	ABSA4832	Ab4832 Human liv
29	33.4	0.7	93	6	ABS22756	Ab22756 Human gen
30	33	0.6	66	2	AAI78908	Aai78908 Poly-glut
31	33	0.6	69	2	AAV17234	Aav17234 SCA2 gene
32	32.4	0.6	87	13	ADQ95148	Adq95148 Synthetic
33	32.2	0.6	69	7	ADS31429	Ads31429 Human gen
34	31.4	0.6	77	4	AAC65410	Aac65410 Oligonucl
35	31.4	0.6	79	4	AAC65411	Aac65411 Oligonucl
36	31.2	0.6	77	13	ADR00028	Adr00028 KRT19 PCR
37	31	0.6	75	12	ADH58952	Adh58952 Silk prot
38	30.8	0.6	99	12	ACH84145	Ach84145 Human gen
39	30.6	0.6	78	2	AAV27572	Aav27572 Nucleotid
40	30.4	0.6	87	13	ADQ95147	Adq95147 Synthetic
41	30.4	0.6	96	6	ABV89298	Abv89298 Human POS
42	30.4	0.6	98	4	AAI84741	Aai84741 Human pol
43	30	0.6	86	3	AAC10510	Aac10510 Human sec
44	29.8	0.6	91	12	ACH89027	Ach89027 Human gen
45	29.4	0.6	57	2	AAx88107	Aax88107 Antifreez

ALIGNMENTS

RESULT 1
ADP49483
ID ADP49483 standard; DNA; 80 BP.
XX
AC ADP49483;
XX
DT 12-AUG-2004 (first entry)
XX
DE Oligonucleotide array related rat oligonucleotide probe No 99.
XX
KW oligonucleotide array; orthologue; homology; expression distribution;
KW change; gene-expression; rat; probe; ss.
XX
OS Rattus norvegicus.
XX
PN JP2004016070-A.
XX
PD 22-JAN-2004.
XX
PF 14-JUN-2002; 2002JP-00174208.
XX
PR 14-JUN-2002; 2002JP-00174208.

(HITA) HITACHI LTD.
WPI; 2004-113862/12.
PT Oligonucleotide array, useful for measuring ortholog gene-expression
PT distribution, comprising number of oligonucleotides derived from ortholog
PT gene of different organism immobilized on support.
XX
XX Disclosure; Page 16; 56pp; Japanese.
XX The invention relates to a novel oligonucleotide array comprising a
XX number of oligonucleotides derived from an orthologue gene of a different
XX organism or species immobilised on a support body. The oligonucleotide
XX array has two sides comprising a gene derived from two different
XX organisms, in which one side comprises a human gene. In the
XX oligonucleotide array a base sequence differs in a different position on
XX the support body. Each of the oligonucleotides is a partial sequence of
XX the orthologue gene of the organism A and organism B. The sequence match
XX degree of the organism species A and the organism species B is less than
XX 70%. The sequence homology of the genes other than the orthologue gene of

the organism species A and the organism species B is the level-of-
statistical-significance value, calculated using the homology searching
algorithm: basic local alignment search tool (BLAST) and the value is 0.1
or more. The found value or the calculated value of the melting-
temperature of the variant genes other than the orthologue gene of the
organism species A and the organism species B is 20 degrees C or more.
The oligonucleotide array is useful for measuring expression distribution
of the orthologue gene in different organisms. The method is useful for
comparing expression change of the orthologue gene with respect to
medical-agent administration and for comparing change of the function of
the orthologue gene in different organisms. The oligonucleotide array has
the ability to perform a measurement of gene-expression distribution of
two or more types of organism simultaneously. The oligonucleotide array
improves reliability of measurement. This polynucleotide sequence
represents a probe of a rat gene for comparison against a human gene used
in the oligonucleotide array of the invention.

Sequence 80 BP; 19 A; 15 C; 25 G; 21 T; 0 U; 0 Other;
Query Match 1.6%; Score 80; DB 12; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.7e-09;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3075 AGACATGTGTTGGAGAGTTCAAGGTCAGAGGGTTGGTTCCCAAGTCTTACGT 3134
1 AGACATGTGTTGGAGAGTTCAAGGTCAGAGGGTTGGTTCCCAAGTCTTACGT 60

3135 GAAACTCATTTGAGGCGCG 3154
1 GAAACTCATTTGAGGCGCG 80

RESULT 2
ABN54476
ID ABN54476 standard; DNA; 65 BP.
XX
AC ABN54476;
XX
DT 15-JUL-2002 (first entry)
XX
XX Mouse spliced transcript detection oligonucleotide SEQ ID NO:27224.
XX
DE Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Mus musculus.
XX
XX WO200210449-A2.
XX
XX 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-IB001903.
XX
XX 28-JUL-2000; 2000US-0221607P.
XX
XX 02-MAY-2001; 2001US-0287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
XX WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
XX selectively hybridize to mRNAs transcribed from a transcription unit of a
XX genome, useful for detecting tissue-, pathology-, and developmental-
XX specific genes.
XX
XX Example 1; SEQ ID NO 27224; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the (sub-
XX)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises several

oligonucleotides, each capable of hybridising selectively to a set of
messenger RNAs transcribed from a given transcription unit of the genome,
which encodes one or more messenger RNA splice variants. The
oligonucleotide libraries are useful for detecting mRNAs from a
biological sample, in expression profiling studies, in qualitatively or
quantitatively characterising the corresponding transcriptome, and in
detecting RNA transcripts and splice variants of human or animal
transcriptomes. The libraries may also be used as specialised mini
libraries to detect transcripts of a sub-transcriptome under a particular
biological or pathological state, and so allowing the detection of tissue
- and pathology-specific genes such as those genes only expressed in
specific tissue under a specific pathological condition; to detect
developmental specific genes; and to detect RNA transcripts and splice
variants of a transcriptome of a patient suffering from a particular
disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
rats, humans and mice, which are used in the exemplification of the
present invention. N.B. the sequence data for this patent did not form
part of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 65 BP; 14 A; 14 C; 17 G; 20 T; 0 U; 0 Other;
Query Match 1.3%; Score 65; DB 6; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4058 GCAGTTTACCTCATTTGACCTTAGTTGCATGTGATCGAAATGTCGAGTCACGTCGA 4117
1 GCAGTTTACCTCATTTGACCTTAGTTGCATGTGATCGAAATGTCGAGTCACGTCGA 60

4118 GAGGC 4122
1 GAGGC 65

RESULT 3
ABN28795
ID ABN28795 standard; DNA; 65 BP.
XX
AC ABN28795;
XX
DT 15-JUL-2002 (first entry)
XX
XX Rat spliced transcript detection oligonucleotide SEQ ID NO:1543.
DE Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Rattus norvegicus.
XX
XX WO200210449-A2.
XX
XX 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-IB001903.
XX
XX 28-JUL-2000; 2000US-0221607P.
XX
XX 02-MAY-2001; 2001US-0287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
XX WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
XX selectively hybridize to mRNAs transcribed from a transcription unit of a
XX genome, useful for detecting tissue-, pathology-, and developmental-
XX specific genes.
XX
XX Example 1; SEQ ID NO 1543; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting

CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-)
 CC transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises several
 CC oligonucleotides, each capable of hybridising selectively to a set of
 CC messenger RNAs transcribed from a given transcription unit of the genome,
 CC which encodes one or more messenger RNA splice variants. The
 CC oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterising the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a particular
 CC biological or pathological state, and so allowing the detection of tissue
 CC - and pathology-specific genes such as those genes only expressed in
 CC specific tissue under a specific pathological condition; to detect
 CC variants of a transcriptome of a patient suffering from a particular
 CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
 CC rats, humans and mice, which are used in the exemplification of the
 CC present invention. N.B. The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 65 BP; 14 A; 13 C; 18 G; 20 T; 0 U; 0 Other;

Query Match 1.1%; Score 54; DB 6; Length 65;
 Best Local Similarity 98.5%; Pred. No. 0.0058;
 Matches 65; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 4056 GAGCAGTTTACCTTCATTTGACCTTACCTTGCATGTGATCGAAATCTCTGAGTCATCGGTG 4115
 Db 1 GAGCAGTTTACCTTCATTTGACCTTACCTTGCATGTGATCGAAATCTCTGAGTC-CTGCGTG 59

QY 4116 CAGAGG 4121
 Db 60 CAGAGG 65

RESULT 4
 ID ABN38018
 ID ABN38018 standard; DNA; 60 BP.
 AC ABN38018;
 XX
 XX 15-JUL-2002 (first entry)
 DT
 XX Human spliced transcript detection oligonucleotide SEQ ID NO:10766.
 DE
 XX Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.
 XX
 XX Homo sapiens.
 OS
 XX WO200210449-A2.
 PN
 XX 07-FEB-2002.
 PD
 XX 20-JUL-2001; 2001WO-IB001903.
 PF
 XX 28-JUL-2000; 2000US-0221607P.
 PR
 XX 02-MAY-2001; 2001US-0287724P.
 XX
 XX (COMP-) COMPUGEN INC.
 PA
 XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
 PI
 XX WPI; 2002-257383/30.
 DR
 XX New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of a
 PT genome, useful for detecting tissue-, pathology-, and developmental-
 PT specific genes.
 XX

PS Example 1; SEQ ID NO 10766; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-)
 CC transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises several
 CC oligonucleotides, each capable of hybridising selectively to a set of
 CC messenger RNAs transcribed from a given transcription unit of the genome,
 CC which encodes one or more messenger RNA splice variants. The
 CC oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterising the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a particular
 CC biological or pathological state, and so allowing the detection of tissue
 CC - and pathology-specific genes such as those genes only expressed in
 CC specific tissue under a specific pathological condition; to detect
 CC variants of a transcriptome of a patient suffering from a particular
 CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
 CC rats, humans and mice, which are used in the exemplification of the
 CC present invention. N.B. The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 60 BP; 19 A; 17 C; 13 G; 11 T; 0 U; 0 Other;

Query Match 1.0%; Score 48.8; DB 6; Length 60;
 Best Local Similarity 88.3%; Pred. No. 0.1;
 Matches 53; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2489 GAGCGCTGTACCCCTTTGAATCCAGAGTCAGATCAGATCCAGTCAGCAGGAGATA 2548
 Db 1 GGGCAGCTGTACCCCTTTGAATCCAGAGTCAGATCAGATCAGTATCCAGCAGGAGACA 60

RESULT 5
 ID ADC16446/c
 ID ADC16446 standard; RNA; 81 BP.
 AC ADC16446;
 XX
 XX 18-DEC-2003 (first entry)
 DT
 XX Short interfering double-stranded RNA oligonucleotide SEQ ID NO:171.
 DE
 XX expression interference; expression inhibition; target gene;
 KW short interfering double stranded RNA; cytostatic; gene therapy;
 KW proliferative disease; cancer; ds.
 XX
 XX Synthetic.
 OS
 XX WO2003012052-A2.
 PN
 XX 13-FEB-2003.
 PD
 XX 30-JUL-2002; 2002WO-US024226.
 PF
 XX 30-JUL-2001; 2001US-0308640P.
 PR
 XX 08-APR-2002; 2002US-0370970P.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (CARN-) CARNEGIE INST WASHINGTON
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 XX Caplen NJ, Morgan RA, Fire A, Parrish S, Mousses S;
 PI Kallionlemi O, Cornelison JR, Alton EW, Griesenbach U;
 XX WPI; 2003-248169/24.
 DR
 XX New RNA comprising double stranded RNA and a 3' or 5' overhang having a
 PT length of 0-nucleotide to 5-nucleotides on each strand, useful as reverse

PT genetic and/or therapeutic tools for interfering or inhibiting expression
 of a target gene.

PS Claim 71; SEQ ID NO 171; 176pp; English.

XX The present invention describes an RNA (I) used for the interference or
 inhibition of expression of a target gene, where (I) comprises double
 stranded RNA of 15-40 nucleotides in length and a 3' or 5' overhang
 having a length of 0-nucleotide to 5-nucleotides on each strand, where
 the sequence of the double stranded RNA is substantially identical to a
 portion of a mRNA or transcript of the target gene. Also described: (1)
 interfering with or inhibiting the expression of a target gene in a cell
 by exposing the cell to an amount of (I); (2) a gene silencing array
 comprising a substantially flat substrate, and addressably arrayed
 different double-stranded RNAs; (3) an array-based method of assessing a
 phenotypic effect of a double-stranded RNA on a target gene; (4)
 validating a gene as a potential drug target for a disease or condition;
 (5) selecting an optimised sequence of a double-stranded RNA for
 interference with or inhibition of expression of a target gene in a cell;
 and (6) a short double-stranded RNA effective for interfering with or
 inhibiting expression of a target gene comprising any of 311 20-78
 nucleotide sequences (see ADC16276 to ADC16586). (I) has cytostatic
 activity, and can be used in gene therapy. The RNAs are useful as reverse
 genetic and/or therapeutic tools for interfering or inhibiting expression
 of a target gene. They are useful for treating proliferative diseases,
 e.g. cancer.

XX Sequence 81 BP; 0 A; 27 C; 27 G; 0 T; 27 U; 0 Other;

Query Match 0.8%; Score 42.8; DB 10; Length 81;
 Best Local Similarity 71.8%; Pred. No. 3.6;
 Matches 56; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1445 AGCGCCAGGAGCAGGAGCGGCAAGCGGAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGC 1504
 |||||
 DB 80 AGC 21
 |||||
 QY 1505 GGGAGCTGGAGCGGCGCAGC 1522
 |||||
 DB 20 AGCAGCAGCAGCAGCAGC 3

RESULT 6

ADC16445

ID ADC16445 standard; RNA; 81 BP.

XX AC ADC16445;

XX 18-DEC-2003 (first entry)

XX Short interfering double-stranded RNA oligonucleotide SEQ ID NO:170.

XX expression interference; expression inhibition; target gene;
 short interfering double stranded RNA; cytostatic; gene therapy;
 proliferative disease; cancer; ds.

XX Synthetic.

XX WO2003012052-A2.

XX 13-FEB-2003.

XX 30-JUL-2002; 2002WO-US024226.

XX 30-JUL-2001; 2001US-0308640P.

XX 08-APR-2002; 2002US-0370970P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 (CARN-) CARNEGIE INST WASHINGTON.
 (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX Caplen NJ, Morgan RA, Fire A, Parrish S, Mousseas S;
 FI Kallioniemi O, Corneliussen JR, Alton EW, Griesenbach U;

XX WPI; 2003-248169/24.

XX New RNA comprising double stranded RNA and a 3' or 5' overhang having a
 length of 0-nucleotide to 5-nucleotides on each strand, useful as reverse
 genetic and/or therapeutic tools for interfering or inhibiting expression
 of a target gene.

PS Claim 71; SEQ ID NO 170; 176pp; English.

XX The present invention describes an RNA (I) used for the interference or
 inhibition of expression of a target gene, where (I) comprises double
 stranded RNA of 15-40 nucleotides in length and a 3' or 5' overhang
 having a length of 0-nucleotide to 5-nucleotides on each strand, where
 the sequence of the double stranded RNA is substantially identical to a
 portion of a mRNA or transcript of the target gene. Also described: (1)
 interfering with or inhibiting the expression of a target gene in a cell
 by exposing the cell to an amount of (I); (2) a gene silencing array
 comprising a substantially flat substrate, and addressably arrayed
 different double-stranded RNAs; (3) an array-based method of assessing a
 phenotypic effect of a double-stranded RNA on a target gene; (4)
 validating a gene as a potential drug target for a disease or condition;
 (5) selecting an optimised sequence of a double-stranded RNA for
 interference with or inhibition of expression of a target gene in a cell;
 and (6) a short double-stranded RNA effective for interfering with or
 inhibiting expression of a target gene comprising any of 311 20-78
 nucleotide sequences (see ADC16276 to ADC16586). (I) has cytostatic
 activity, and can be used in gene therapy. The RNAs are useful as reverse
 genetic and/or therapeutic tools for interfering or inhibiting expression
 of a target gene. They are useful for treating proliferative diseases,
 e.g. cancer.

XX Sequence 81 BP; 27 A; 27 C; 27 G; 0 T; 0 U; 0 Other;

Query Match 0.8%; Score 42.8; DB 10; Length 81;
 Best Local Similarity 71.8%; Pred. No. 3.6;
 Matches 56; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1445 AGCGCCAGGAGCAGGAGCGGCAAGCGGAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGC 1504
 |||||
 DB 2 AGC 61
 |||||

QY 1505 GGGAGCTGGAGCGGCGCAGC 1522
 |||||

DB 62 AGC 79

RESULT 7

AAC65404/c

ID AAC65404 standard; DNA; 98 BP.

XX AAC65404;

XX 14-FEB-2001 (first entry)

XX Oligonucleotide B6463 for chimeric receptor construction.

XX Membrane-associated protein; antiviral; antibacterial; antiparasitic;
 immunomodulatory; anticancer; antiinflammatory; antiasthmatic;
 antidiabetic; neuroprotective; chimeric receptor; infection;
 inflammatory disorder; autoimmune disorder; cancer; allergy; asthma;
 eczema; cystic fibrosis; sickle cell anaemia; psoriasis;
 multiple sclerosis; organ transplant rejection; diabetes;
 transmembrane domain; ss.

XX Synthetic.

XX WO200063373-A1.

XX 26-OCT-2000.

XX 17-APR-2000; 2000WO-GB001471.

XX Lahue RS, Pelletier R, Miret JJ;
 XX WPI; 2002-575460/61.
 XX
 XX Detecting alterations in trinucleotide repeat (TNR) tract lengths in
 XX mammalian cells useful for identifying disorders associated with TNR
 XX instability such as cancer, Fragile X syndrome, Huntington's disease or
 XX myotonic dystrophy.
 XX
 XX Example 1; Page 24; 43pp; English.
 XX
 XX The invention relates to detecting (M1) alterations in trinucleotide
 XX repeat (TNR) tract lengths comprising: (i) contacting mammalian cells
 XX with a shuttle vector (containing approximately 25 repeats) that enters
 XX and replicates in the cells; (ii) recovering the vector into a yeast cell in
 XX the presence of a selection agent or (where the expanded TNR confers a
 XX His⁺ phenotype) in the absence of histidine, alteration to the TNR tract
 XX confers resistance to the selective agent or His⁺ phenotype; and (iv)
 XX selecting yeast cells with the shuttle vector containing TNR tract
 XX alterations that survived the presence of the selective agent. Also
 XX included is an adaptation of the used to detect the contraction of 33 or
 XX 55 repeats using growth of the yeast cells in the absence of uracil. The
 XX methods of the present invention are useful for identifying disorders
 XX with genetic alterations associated with TNR instability such as cancer
 XX (e.g. testicular and prostate), Fragile X syndrome, Huntington's disease,
 XX myotonic dystrophy, spinal and bulbar muscular atrophy (SBMA),
 XX spinocerebellar ataxias (types 1, 8, 12, 3 (Machado-Joseph disease), 6,
 XX 7 and 2), dentatorubral-pallidoluysian atrophy (DRPLA), Friedreich's
 XX ataxia and Kennedy's disease. The present sequence is the CAG25 TNR known
 XX to be unstable in Human sperm and in yeast which was tested in the method
 XX of the invention for expansion
 XX
 XX SQ Sequence 75 BP; 25 A; 25 C; 25 G; 0 T; 0 U; 0 Other;

Query Match 0.8%; Score 41; DB 6; Length 75;
 Best Local Similarity 72.6%; Pred. No. 9.5;
 Matches 53; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
 QY 1450 CAGGAGCAGGCGCAGGCGGAGCTGGAGTGGAGAGCAGCTGGAGAGCAGCGGAG 1509
 Db 1 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 60
 QY 1510 CTGAGCGGCGAGC 1522
 Db 61 CAGCAGCAGCAGC 73

RESULT 10
 AAC65406/c
 ID AAC65406 standard; DNA; 92 BP.
 XX AAC65406;
 XX
 XX 14-FEB-2001 (first entry)
 XX
 XX Oligonucleotide B6465 for chimeric receptor construction.
 XX Membrane-associated protein; antiviral; antibacterial; antiparasitic;
 XX immunomodulatory; anticancer; antiinflammatory; antiasthmatic;
 XX antidiabetic; neuroprotective; chimeric receptor; infection;
 XX inflammatory disorder; autoimmune disorder; cancer; allergy; asthma;
 XX eczema; cystic fibrosis; sickle cell anaemia; psoriasis;
 XX multiple sclerosis; organ transplant rejection; diabetes;
 XX transmembrane domain; ss.
 XX Synthetic.
 XX OS
 XX WO200063373-A1.
 XX
 XX 26-OCT-2000.
 XX
 XX 17-APR-2000; 2000WO-GB001471.

PF 17-APR-2000; 2000WO-GB001471.
 XX
 XX 16-APR-1999; 99GB-00008816.
 XX
 XX (CLLT) CELLTECH THERAPEUTICS LTD.
 XX Finney HM, Lawson ADG;
 XX WPI; 2001-015774/02.
 XX
 XX Altering the properties or level of expression of membrane-associated
 XX proteins, e.g., to change responses to cell surface antigens or the
 XX sensitivity of intracellular signaling.
 XX
 XX Example 2; Fig 3; 47pp; English.
 XX
 XX The present sequence was used in the construction of chimeric receptors.
 XX The properties and level of expression of a membrane-associated protein
 XX may be altered by substituting a transmembrane region or a membrane-
 XX anchoring region for transmembrane or membrane-anchoring regions that are
 XX not naturally part of the protein. The relative response of membrane-
 XX associated proteins to cell surface-associated antigen versus antigen in
 XX solution, and the sensitivity of intracellular signaling mediated by
 XX membrane-associated proteins can be altered. They can be important in
 XX treatment of, e.g. HIV infection, bacterial infections, parasitic
 XX infections, inflammatory/autoimmune disorders (e.g. rheumatoid arthritis,
 XX osteoarthritis or inflammatory bowel disease), cancer, allergic/atopic
 XX diseases (e.g. asthma or eczema), congenital disorders (e.g. cystic
 XX fibrosis or sickle cell anaemia), dermatological disorders (e.g.
 XX psoriasis), neurological disorders (e.g. multiple sclerosis), organ
 XX transplant rejection, graft-versus-host diseases, or metabolic/idiopathic
 XX diseases (e.g. diabetes)
 XX
 XX SQ Sequence 92 BP; 2 A; 34 C; 24 G; 32 T; 0 U; 0 Other;

Query Match 0.8%; Score 40.4; DB 4; Length 92;
 Best Local Similarity 71.8%; Pred. No. 15;
 Matches 53; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
 QY 1437 AGAGCGGAGCGCGCAGGAGCAGGAGCGGAGCGGAGCTGGAGTGGAGAGCAGCTGGA 1496
 Db 86 AGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 27
 QY 1497 GAAGCAGCGGAGC 1510
 Db 26 GCAGGAGCAGCAGC 13

RESULT 11
 AAC65407
 ID AAC65407 standard; DNA; 94 BP.
 XX AAC65407;
 XX
 XX 14-FEB-2001 (first entry)
 XX
 XX Oligonucleotide B6466 for chimeric receptor construction.
 XX Membrane-associated protein; antiviral; antibacterial; antiparasitic;
 XX immunomodulatory; anticancer; antiinflammatory; antiasthmatic;
 XX antidiabetic; neuroprotective; chimeric receptor; infection;
 XX inflammatory disorder; autoimmune disorder; cancer; allergy; asthma;
 XX eczema; cystic fibrosis; sickle cell anaemia; psoriasis;
 XX multiple sclerosis; organ transplant rejection; diabetes;
 XX transmembrane domain; ss.
 XX Synthetic.
 XX OS
 XX WO200063373-A1.
 XX
 XX 26-OCT-2000.
 XX
 XX 17-APR-2000; 2000WO-GB001471.

DE Clone #8 fragment identified by CAG repeat analysis method.
 XX CAG repeat; human; genome analysis; medical diagnostic;
 KW nucleic acid analysis; variation assessment; neurological disease;
 KW Huntington's chorea; PCR suppression; ss.
 XX Homo sapiens.
 XX
 XX WO9849345-A1.
 XX
 XX 05-NOV-1998.
 XX
 XX 29-APR-1998; 98WO-US008616.
 XX
 XX 29-APR-1997; 97US-0045078P.
 XX
 XX (UYBO-) UNIV BOSTON.
 XX
 XX Smith CL;
 XX
 XX WPI; 1998-594983/50.
 XX
 XX Analysing nucleic acid samples - using amplification primers which
 PT contain CAG or CTG tri-nucleotide repeats for differential display of
 PT samples from different sources.
 XX
 XX Example; Page 32; 44pp; English.
 XX
 XX This sequence represents a fragment of a human CAG repeat containing
 CC clone DNA sequence isolated using the method of the invention. The method
 CC is for analysing nucleic acids in a sample, and comprises: (a) providing
 CC a sample containing nucleic acid, a first oligonucleotide primer
 CC comprising a CTG repeat, a second oligonucleotide primer comprising a CAG
 CC repeat and a polymerase and PCR reagents; (b) preparing the nucleic acid
 CC so that it is amplifiable; (c) amplifying the nucleic acid with the first
 CC and second primers; and (d) detecting the amplified product. The method
 CC is used to distinguish between the expression of genes in two or more
 CC biological samples, e.g. body fluids, cells, solid tissue or solid and
 CC liquid foods. It can be used in medical diagnostics, e.g. to
 CC differentiate between normal and diseased tissue or to assess the
 CC variation within monozygotic twin pairs. The method allows the isolation
 CC and analysis of genome subsets containing CAG repeats which are known to
 CC be important in a number of neurological diseases including Huntington's
 CC chorea. The method uses PCR suppression, in which only fragments which
 CC contain a target repeat are efficiently amplified. This allows accurate
 CC identification of differentially expressed genes in various cell types.
 CC Genome complexity is reduced by the new method which targets genomic
 CC subsets containing CAG repeats
 XX
 XX Sequence 89 BP; 2 A; 23 C; 37 G; 27 T; 0 U; 0 Other;
 SQ
 Query Match 0.8%; Score 39; DB 2; Length 89;
 Best Local Similarity 65.5%; Pred. No. 32;
 Matches 57; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
 QY 1441 CGGAGCGCCAGCAGCAGGAGCCAGCGGAGCTGGAGCTGGAGAGCAGCTGGAGAG 1500
 DB 89 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 30
 QY 1501 CAGCGGAGCTGGAGCGGCGCAGCGAG 1527
 DB 29 CAGCCCCCAGCAGCAGCAGCAGCAG 3
 RESULT 14
 AAV17232
 ID AAV17232 standard; DNA; 78 BP.
 XX
 XX AAV17232;
 XX
 XX 29-JUN-1998 (first entry)
 DT
 XX SCA2 gene CAG repeat unit fragment.
 XX SCA2 gene; spinocerebellar ataxis type II; CAG repeat; PCR primer; ss.
 XX Synthetic.
 XX WO9803679-A1.
 XX
 XX 29-JAN-1998.
 XX
 XX 18-JUL-1996; 96WO-JP001999.
 XX
 XX 18-JUL-1996; 96WO-JP001999.
 XX
 XX (SRLS-) SRL INC.
 XX
 XX Tsuji S, Sanpei K;
 XX
 XX WPI; 1998-120796/11.
 DE

XX SCA2 gene; spinocerebellar ataxis type II; CAG repeat; PCR primer; ss.
 KW Synthetic.
 XX WO9803679-A1.
 XX
 XX 29-JAN-1998.
 XX
 XX 18-JUL-1996; 96WO-JP001999.
 XX
 XX 18-JUL-1996; 96WO-JP001999.
 XX
 XX (SRLS-) SRL INC.
 XX
 XX Tsuji S, Sanpei K;
 XX
 XX WPI; 1998-120796/11.
 DE

XX Sequence 78 BP; 23 A; 30 C; 25 G; 0 T; 0 U; 0 Other;
 SQ
 Query Match 0.7%; Score 38; DB 2; Length 78;
 Best Local Similarity 67.3%; Pred. No. 53;
 Matches 53; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
 QY 1450 CAGCAGCAGGAGCCAGCGGAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAG 1509
 DB 1 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 60
 QY 1510 CTGAGCGGCGCAGCAGAG 1527
 DB 61 CAGCAGCGCGCGCCGCG 78
 RESULT 15
 AAV17233
 ID AAV17233 standard; DNA; 78 BP.
 XX
 XX AAV17233;
 XX
 XX 29-JUN-1998 (first entry)
 DT
 XX SCA2 gene CAG repeat unit fragment.
 XX SCA2 gene; spinocerebellar ataxis type II; CAG repeat; PCR primer; ss.
 XX Synthetic.
 XX WO9803679-A1.
 XX
 XX 29-JAN-1998.
 XX
 XX 18-JUL-1996; 96WO-JP001999.
 XX
 XX 18-JUL-1996; 96WO-JP001999.
 XX
 XX (SRLS-) SRL INC.
 XX
 XX Tsuji S, Sanpei K;
 XX
 XX WPI; 1998-120796/11.
 DE

This sequence represents a fragment of the SCA2 gene. It can be used in the method of the invention for diagnosing spinocerebellar ataxis type II, by performing PCR on the test DNA using two primers hybridising to parts of the SCA2 gene sequence, and determining the number of CAG repeats in the amplified products. The method provides an easy means for the diagnosis of spinocerebellar ataxis type II

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 06:48:09 ; Search time 15136.1 Seconds
(without alignments)
12785.281 Million cell updates/sec

Title: US-09-674-237B-1
Perfect score: 5084
Sequence: 1 cggcagaggaggaggag.....gagaattgatcatcaagctt 5084

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 675282

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	1.8	94	8	CC199586 XE464 Bay
2	74.2	1.5	79	1	AA589590 v149c09.s
3	54.6	1.1	64	9	HSNC18D03
4	46	0.9	46	8	CC200157 XG352 Bay
5	44	0.9	92	8	AZ786638 2M0032E18
6	42.8	0.8	89	8	AZ448190 1M0245A17
7	40.6	0.8	86	1	AV533640 AV533640
8	40.4	0.8	94	8	AZ804535 2M0065I21
9	36.4	0.7	71	8	AZ471347 1M0285I21
10	36.2	0.7	87	8	AZ837486 2M0132J14
11	36.2	0.7	91	8	AZ614077 1M0442B14
12	36	0.7	91	8	BH225497 1006126F0
13	36	0.7	91	8	BH225515 1006126G0
14	36	0.7	91	8	BH225549 1006126H0
15	35.6	0.7	96	9	CR272899 Forward s
16	35.4	0.7	89	9	CR256426 Reverse s
17	34.2	0.7	95	8	AZ779622 2M0016I14
18	34.2	0.7	98	8	AQ988931 26A1A04NE
19	34	0.7	84	9	CG565168 OST189766
20	34	0.7	92	6	CA340475 NISC 1z04
21	34	0.7	96	7	CV030123 9138-Full
22	33.8	0.7	80	8	AZ408358 1M0179F06
23	33.6	0.7	83	7	CO755195 Mdfit3048
24	33.2	0.7	100	7	CK459729 929944 MA

25 33 CG572531 84 9 CG572531
26 32.8 BH218452 98 8 BH218452
27 32.4 CD487616 98 6 CD487616
c 28 32.2 AA726119 91 1 AA726119
29 32 CR126087 60 9 CR126087
30 31.8 CG572499 87 9 CG572499
c 31 31.8 AZ476761 94 8 AZ476761
32 31.8 BQ234425 100 5 BQ234425
33 31.6 CR147383 86 9 CR147383
34 31.6 BQ809943 88 5 BQ809943
35 31.6 AI316382 91 1 AI316382
36 31.6 CV295497 100 7 CV295497
c 37 31.4 DI8610 100 7 DI8610
38 31.2 CG724316 85 9 CG724316
39 31.2 BE662032 88 2 BE662032
40 31.2 BQ234381 95 5 BQ234381
41 31.2 BQ234483 96 5 BQ234483
42 31.2 AI947161 97 1 AI947161
43 31.2 AV674805 98 1 AV674805
c 44 31.2 BQ237998 100 5 BQ237998
45 31 BG393282 79 4 BG393282

ALIGNMENTS

RESULT 1
LOCUS CC199586 94 bp mRNA linear GSS 09-MAY-2003
DEFINITION XE464 BayGenomics Gene Trap Library pGTLXf Mus musculus cDNA, mRNA
sequence.
ACCESSION CC199586
VERSION CC199586.1 GI:30479626
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 94)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS BayGenomics.
TITLE http://baygenomics.ucsf.edu/
JOURNAL Unpublished (2001)
COMMENT Contact: BayGenomics
Bay Area Functional Genomics Consortium (BayGenomics)
Email: info@baygenomics.ucsf.edu
Sequence tag generated by 5' RACE of total RNA from gene trap ES
cell line. ES cell lines harboring insertion mutation of target
gene are available upon request from BayGenomics. Annotation
information available from
http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=
CELL LINE&KEY=XE464
Class: Gene Trap.
Location/Qualifiers
1. .94
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129.ola"
/db_xref="taxon:10090"
/sex="Male"
/clone_lib="Embryonic stem cell"
/clone_type="BayGenomics Gene Trap Library pGTLXf"
/note="Vector: pGTLXf"

FEATURES

source
1. .94
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129.ola"
/db_xref="taxon:10090"
/sex="Male"
/clone_lib="Embryonic stem cell"
/clone_type="BayGenomics Gene Trap Library pGTLXf"
/note="Vector: pGTLXf"

ORIGIN

Query Match 1.8%; Score 94; DB 8; Length 94;
Best Local Similarity 100.0%; Pred. No. 2.2e-12;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 886 CCTCCAGCAGCAGATGGCTGCTCAGTCATCAAGCTGCAATACAGGAGTTATTC 945
Db 1 CCTCCAGCAGCAGATGGCTGCTCAGTCATCAAGCTGCAATACAGGAGTTATTC 60
QY 946 AACAGCCACGACAAACTATGATGACACTTAA 979

```

|||||
61 AACAGCCACGACAAACTATGAGTGGACACTTAA 94

RESULT 2
AA589590 79 bp mRNA linear EST 16-SEP-1997
LOCUS v149c09.s1 Stragatene mouse skin (#937313) Mus musculus cDNA clone
DEFINITION IMAGE:975568 3', similar to SW:YFJ4 YEAST P43603 HYPOTHETICAL 40.4
KD PROTEIN IN PES4-HIS2 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION AA589590
VERSION AA589590.1 GI:2402970
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 79)
Marrá,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:556296
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 1.
FEATURES
source
1..79
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:975568"
/sex="females"
/tissue_type="whole skin"
/dev_stage="11 weeks old"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse skin (#937313)"
/notes="Organ: skin; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGGCAGG 3' -3' adaptor
sequence: 5' CTCGAGTGTGTTTTTTTTTTT 3'"
ORIGIN
Query Match 1.58; Score 74.2; DB 1; Length 79;
Best Local Similarity 96.24; Pred. No. 2.3e-07;
Matches 76; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3261 TGCATGTACATACAGAGAGTCTGTGACGAGGAGATTTAACTTTTCAGCAAGGGATGT 3320
Db 1 TGCATGTACATACAGAGAGTCTGTGATCAAGGAGATTAAAGTTTCAGCAAGGGATGT 60
QY 3321 GATTGTGTTACCAAGAAA 3339
Db 61 GATTGTGTTACCAAGAAA 79
RESULT 3
HSMC18D03
LOCUS HSMC18D03 64 bp DNA linear GSS 29-MAY-1997
DEFINITION H.sapiens DNA for trapped exon (ID HMC18D03), genomic survey
sequence.
ACCESSION X88324
VERSION X88324.1 GI:1437729
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 64)
Chen,H., Chrast,R., Rossier,C., Morris,M.A., Lalioti,M.D. and
Antonarakis,S.E.
Cloning of 559 potential exons of genes of human chromosome 21 by
exon trapping
JOURNAL Genome Res. 6 (8), 747-760 (1996)
MEDLINE 97011340
PubMed 8858350
REFERENCE 2 (bases 1 to 64)
AUTHORS Chen,H.M., Rossier,C., Chrast,R. and Antonarakis,S.E.
TITLE Cloning of trapped exons from human chromosome 21
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 64)
AUTHORS Antonarakis,S.E.
TITLE Direct Submission
JOURNAL Submitted (17-MAR-1995) Stylianos E. Antonarakis, Division of
Medical Genetics, University and Cantonal Hospital of Geneva, CMU,
1 rue Michel-Servet, 1211 Geneva, SWITZERLAND
FEATURES
source
1..64
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
1..64
/feature="trapped exon"
ORIGIN
Query Match 1.11; Score 54.6; DB 9; Length 64;
Best Local Similarity 89.11; Pred. No. 0.021;
Matches 57; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 983 GTCCCCAGGCAAGAACTATTCATGCAATCAAGTTTACCCAGGCTCAGCTGGCTTCAA 1042
Db 1 GTCCCCAGGCAAGAACTATTCATGCAATCAAGTTTACCCAGGCTCAGCTGGCTTCAA 60
QY 1043 TATG 1046
Db 61 TATG 64
RESULT 4
CC200157
LOCUS CC200157
DEFINITION XG352 BayGenomics Gene Trap Library pGTLXf Mus musculus cDNA, mRNA
sequence.
ACCESSION CC200157
VERSION CC200157.1 GI:30479920
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 46)
http://baygenomics.ucsf.edu/
JOURNAL Unpublished (2001)
COMMENT Contact: BayGenomics
Bay Area Functional Genomics Consortium (BayGenomics)
Email: info@baygenomics.ucsf.edu
Sequence tag generated by 5' RACE of total RNA from gene trap ES
cell line. ES cell lines harboring insertion mutation of target
gene are available upon request from BayGenomics. Annotation

```

information available from
http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL_LINE&KEY=XG352
 Class: Gene Trap.

FEATURES

source
 1..46
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129 ola"
 /db_xref="taxon:10090"
 /sex="Male"
 /cell_type="Embryonic stem cell"
 /clone_lib="BayGenomics Gene Trap Library pGTLxf"
 /notes="Vector: pGTLxf"

ORIGIN

Query Match 0.9%; Score 46; DB 8; Length 46;
 Best Local Similarity 100.0%; Pred. No. 3;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 558 TGTCATGAACACCAACCACTGCTATTTCCAGTGCACCACTTTT 603
 |||||
 Db 1 TGTCATGAACACCAACCACTGCTATTTCCAGTGCACCACTTT 46

RESULT 5

AZ786638/c
 LOCUS 92 bp DNA linear GSS 16-FEB-2001
 DEFINITION 2M0032E18F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC2M0032E18 F, genomic survey sequence.

ACCESSION AZ786638
 VERSION AZ786638.1 GI:12924599
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS 1 (bases 1 to 92)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
 Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
 Niederhausern, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

JOURNAL

COMMENT Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA

Tel: 801 585 5606
 Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00

Plate: 0032 row: E column: 18

Seq primer: CGTTGTAACAGCAGCCAGT

Class: plasmid ends

High quality sequence stop: 92.

FEATURES

source

1..92
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0032E18"
 /sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PWD42 (GI:4732114|GB|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

ORIGIN

Query Match 0.9%; Score 44; DB 8; Length 92;
 Best Local Similarity 67.4%; Pred. No. 12;
 Matches 62; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1454 AGCAGGAGCCCAAGCGCGAGCTGGAGAGAGAGAGAGAGAGAGAG 1513
 |||||
 Db 92 AGGAGGAGGAGTAGGAGTAGAGGAGAGAGAGAGAGAGAGAGAG 33

QY 1514 AGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1545
 |||||

Db 32 AGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1

RESULT 6

AZ448190

LOCUS

DEFINITION

89 bp DNA linear GSS 04-OCT-2000
 clone UUGC1M0245A17 R, genomic survey sequence.

ACCESSION AZ448190

VERSION AZ448190.1 GI:10600743

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS 1 (bases 1 to 89)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
 Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
 Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0245 row: A column: 17

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 89.

Location/Qualifiers

1..89

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0245A17"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/note="Vector: PWD42nv; Purified genomic DNA from M.

was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 [gi|4732114|gb|AF129072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

[illegible]

RESULT 12	
BH225497/c	
LOCUS	
DEFINITION	1006126508.x1 1006 - RescueMu Grid G Zea mays genomic, genomic survey sequence.
ACCESSION	BH225497
VERSION	BH225497.1
KEYWORDS	GI:16823515
SOURCE	GSS.
ORGANISM	Zea mays
	Zea mays
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE	1. (bases 1 to 91)
AUTHORS	Walbot, V.
TITLE	Maize genomic sequences found using engineered RescueMu transposon
JOURNAL	Unpublished (2001)
COMMENT	Contact: Walbot, V.

Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site so sequence was trimmed. Post-ligation
sequence submitted separately.
Plate: 1006126 row: 1
Class: transposon-tagged.

FEATURES	source
Location/Qualifiers	
1. .91	
/organism="Zea mays"	
/mol_type="genomic DNA"	
/cultivar="mixed background W23/A188/B73"	
/db xref="taxon:4577"	
/tissue_type="leaf"	
/dev_stage="adult"	
/lab_host="DH10B"	
/clone_lib="1006 - RescueMu Grid G"	
/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmmb.iastate.edu' and follow the links for	

10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (G1473211[gb]AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

[illegible][illegible]

TITLE	Mouse whole genome scaffolding with paired end reads from 10x
JOURNAL	plasmid inserts
COMMENT	Unpublished (2000) Contact: Robert B. Weiss University of Utah Genome Center University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0442 row: B column: 14 Seq primer: CACACAGGAACAGCTATGACC Class: plasmid ends Unit: base pair

```

FEATURES
source
high quality sequence stop: 21.
Location/Qualifiers
1..91
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC1M0442B14"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
needle at constant velocity. The sheared DNA

```

QY 1434 GAAAGAGCGGGAGCGCCAGGAGCA 1457

Db 25 GAAGACCCAGGACCGCGCAGCA 2

RESULT 15

CR272899/c

LOCUS

DEFINITION

Forward strand read from insert in 5'HPRT insertion targeting and

chromosome engineering clone MHPN79102, genomic survey sequence.

ACCESSION

CR272899

VERSION

GSS; genome survey sequence; MICER.

KEYWORDS

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 96)

Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,

Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,

Rogers,J. and Bradley,A.

Direct Submission

Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>

FEATURES

Location/Qualifiers

1..96

source

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="MHPN79102"

/clone_lib="MHPN"

ORIGIN

Query Match 0.7%; Score 35.6; DB 9; Length 96;

Best Local Similarity 62.2%; Pred.No.1.7e+03;

Matches 56; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 1465 AAGCGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAGCGGAGCGGA 1524

Db 95 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGG 36

QY 1525 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1554

Db 35 AAGAGGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGG 6

Search completed: April 29, 2005, 02:01:57
Job time : 15142.1 secs

Result No.	Score	Query #		Length	DB	ID	Description
		Match					
1	38	0.7	78	3	US-09-043-303-12		Sequence 12, Appl
2	37.6	0.7	69	4	US-09-573-080A-463		Sequence 463, App
3	36.4	0.7	78	3	US-09-043-303-11		Sequence 11, Appl
4	34	0.7	75	3	US-09-043-303-10		Sequence 10, Appl
5	33	0.6	69	3	US-09-043-303-13		Sequence 13, Appl
6	32.2	0.6	69	4	US-09-573-080A-462		Sequence 462, App
C 7	31.4	0.6	89	4	US-09-270-767-4923		Sequence 4823, A
C 8	31.4	0.6	89	4	US-09-270-767-20105		Sequence 20105, A
9	30.4	0.6	93	3	US-08-556-978B-25		Sequence 24, Appl
C 10	30.4	0.6	93	3	US-08-556-978B-25		Sequence 25, Appl
C 11	30.2	0.6	51	1	US-08-068-747-1		Sequence 1, Appl
12	30	0.6	86	4	US-09-513-999C-14585		Sequence 14585, A
13	29.4	0.6	57	2	US-07-814-220-26		Sequence 26, Appl
14	29.4	0.6	57	2	US-07-812-421-26		Sequence 26, Appl
C 15	29.4	0.6	62	2	US-07-814-220-25		Sequence 25, Appl
C 16	29.4	0.6	62	2	US-07-812-421-25		Sequence 25, Appl
17	29.4	0.6	97	1	US-08-182-175A-48		Sequence 48, Appl
18	29.4	0.6	97	1	US-08-474-633A-57		Sequence 57, Appl
19	29.4	0.6	97	3	US-08-823-771-57		Sequence 57, Appl
20	29.4	0.6	97	5	PCF-US92-06412-48		Sequence 48, Appl
C 21	29.2	0.6	99	4	US-09-402-532-21		Sequence 21, Appl
C 22	29	0.6	96	3	US-09-281-481A-6		Sequence 6, Appl
C 23	29	0.6	96	3	US-09-281-481A-7		Sequence 7, Appl
24	28.6	0.6	94	4	US-09-513-999C-35527		Sequence 3527, A
25	28.4	0.6	94	4	US-09-513-999C-29269		Sequence 29269, A
26	28.2	0.6	57	3	US-09-043-303-14		Sequence 14, Appl
27	28.2	0.6	58	3	US-08-860-038-15		Sequence 15, Appl

```

; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: repetitive sequence found in
; OTHER INFORMATION: many eutherian genomes. Length of core repeating element is vari
; OTHER INFORMATION: able and is often polymorphic
US-09-573-080A-463

Query Match      0.7%; Score 37.6; DB 4; Length 69;
Best Local Similarity 72.1%; Pred. No. 5;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1455 GCAGGAGCCAGCGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGGAGCTGGA 1514
DB 1 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 60

QY 1515 GCGGCAGC 1522
DB 61 GCAGCAGC 68

RESULT 3
US-09-043-303-11
; Sequence 11, Application US/09043303
; Patent No. 6251589
; GENERAL INFORMATION:
; APPLICANT: TSUJI, Shoji
; APPLICANT: SANPEI, Kazujiro
; TITLE OF INVENTION: Method for Diagnosing Spinocerebellar Alaxia Type 2 and
; TITLE OF INVENTION: Primers Therefor
; FILE REFERENCE: 0760-0241P
; CURRENT APPLICATION NUMBER: US/09/043,303
; CURRENT FILING DATE: 1998-05-18
; EARLIER APPLICATION NUMBER: PCT/JP96/01999
; EARLIER FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 78
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-043-303-11

Query Match      0.7%; Score 36.4; DB 3; Length 78;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 52; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1450 CAGCAGCAGCAGCGCAAGCGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGGAG 1509
DB 1 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 60

QY 1510 CTGAGCGCGCAGCAGAG 1527
DB 61 CAGCAGCGCGCGCCGCG 78

RESULT 4
US-09-043-303-10
; Sequence 10, Application US/09043303
; Patent No. 6251589
; GENERAL INFORMATION:
; APPLICANT: TSUJI, Shoji
; APPLICANT: SANPEI, Kazujiro
; TITLE OF INVENTION: Method for Diagnosing Spinocerebellar Alaxia Type 2 and
; TITLE OF INVENTION: Primers Therefor
; FILE REFERENCE: 0760-0241P
; CURRENT APPLICATION NUMBER: US/09/043,303
; CURRENT FILING DATE: 1998-05-18
; EARLIER APPLICATION NUMBER: PCT/JP96/01999
; EARLIER FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10

```

```

; LENGTH: 75
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-043-303-10

Query Match      0.7%; Score 34; DB 3; Length 75;
Best Local Similarity 66.2%; Pred. No. 47;
Matches 49; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1454 ACAGGAGCCAGCGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGGAGCTGG 1513
DB 2 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 61

QY 1514 AGCGCAGCAGAG 1527
DB 62 AGCGCGCGCCGCG 75

RESULT 5
US-09-043-303-13
; Sequence 13, Application US/09043303
; Patent No. 6251589
; GENERAL INFORMATION:
; APPLICANT: TSUJI, Shoji
; APPLICANT: SANPEI, Kazujiro
; TITLE OF INVENTION: Method for Diagnosing Spinocerebellar Alaxia Type 2 and
; TITLE OF INVENTION: Primers Therefor
; FILE REFERENCE: 0760-0241P
; CURRENT APPLICATION NUMBER: US/09/043,303
; CURRENT FILING DATE: 1998-05-18
; EARLIER APPLICATION NUMBER: PCT/JP96/01999
; EARLIER FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 69
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-043-303-13

Query Match      0.6%; Score 33; DB 3; Length 69;
Best Local Similarity 73.7%; Pred. No. 83;
Matches 42; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1466 AGCGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGGAGCTGGAGCGGCAGC 1522
DB 2 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 58

RESULT 6
US-09-573-080A-462
; Sequence 462, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 462
; LENGTH: 69
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: repetitive sequence found in
; OTHER INFORMATION: many eutherian genomes. Length of core repeating element is varia
; OTHER INFORMATION: ble and is often polymorphic
US-09-573-080A-462

Query Match      0.6%; Score 32.2; DB 4; Length 69;

```

Query Match	0.6%	Score 31.4;	DB 4;	Length 89;
Best Local Similarity	61.7%	Pred. No. 2.6e+02;		
Matches 50; Conservative	0;	Mismatches 31;	Indels 0;	Gaps 0;
Qy	1467	GCGCACCTGAGCTGAGAAAGCAGCTGGAGAACACACGGGAGCTGGAGCGGAGCGGAGG	1526	
Db	81	GCGCCCTGAGCTGAGCAGGAGCGCAGATCACACCGGCTGCTGGAGCAGGAAGGAC	22	
Qy	1527	GGAGGAGGAGGAAGGAGAT	1547	
Db	21	TGCTGAATCGAGCTGGAGCT	1	

```

RESULT 10
US-08-556-978B-25/c
; Sequence 25, Application US/08556978B
; Patent No. 6268169
;
; GENERAL INFORMATION:
;
; APPLICANT: FAHNESTOCK, STEPHEN F.
;
; TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
;
; TITLE OF INVENTION: SPIDER SILK ANALOGS
;
; NUMBER OF SEQUENCES: 107
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSES: E. I. DU PONT DE NEMOURS AND COMPANY
;
; STREET: 1007 MARKET STREET
;
; CITY: WILMINTON
;
; STATE: DELAWARE
;
; COUNTRY: UNITED STATES OF AMERICA
;
; ZIP: 19898
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: DISKETTE, 3.50 INCH
;
; COMPUTER: IBM PC COMPATIBLE

```

```

; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556,978B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: 08/077,600
; FILING DATE: JUNE 15, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9389-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-556-978B-25

Query Match 0.6%; Score 30.4; DB 3; Length 93;
Best Local Similarity 71.4%; Pred. No. 4.9e+02;
Matches 40; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1966 GCCTTGAAGCAAGAGCTGGCCCGGACAGCTCCGGGAGCAGCTGGACGAGGT 2021
Db 75 GCCTTGTGTCAGGCTGCTGGCCGGACAGCTGGTGGTGGTGGT 20

RESULT 11
US-08-674-747-1/c
; Sequence 1, Application US/08068747
; Patent No. 5695933
; GENERAL INFORMATION:
; APPLICANT: Schalling, Martin
; APPLICANT: Hudson, Thomas J.
; APPLICANT: Housman, David E.
; TITLE OF INVENTION: Direct Determination of Expanded
; TITLE OF INVENTION: Nucleotide Repeats in the Human Genome
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,747
; FILING DATE: 28-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-6141
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

us-09-674-237b-1.szlm100.rni

; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic"
; US-08-068-747-1

Query Match 0.6%; Score 30.2; DB 1; Length 51;
Best Local Similarity 74.5%; Pred. No. 3.7e+02;
Matches 38; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1471 CAGCTGAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAGCGGCGAG 1521
Db 51 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1

RESULT 12
US-09-513-999C-14585
; Sequence 14585, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14585
; LENGTH: 86
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 40
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 57
; OTHER INFORMATION: r=a or g
; US-09-513-999C-14585

Query Match 0.6%; Score 30; DB 4; Length 86;
Best Local Similarity 59.3%; Pred. No. 6e+02;
Matches 48; Conservative 1; Mismatches 32; Indels 0; Gaps 0;

QY 1465 AAGCGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAGCGGCGGA 1524
Db 4 AAGCGCGCTCGCGCGCTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCGGCG 63

QY 1525 GAGGAGGAGGAGGAGGAGGAG 1545
Db 64 GCGATGCTGAAGATGGCGCG 94

RESULT 13
US-07-814-220-26
; Sequence 26, Application US/07814220
; Patent No. 5925540
; GENERAL INFORMATION:
; APPLICANT: Caceci, Thomas
; APPLICANT: Toth, Thomas E.
; APPLICANT: Szumanski, Maria B.W.
; TITLE OF INVENTION: SYNTHETIC ANTIFREEZE PEPTIDE AND
; TITLE OF INVENTION: SYNTHETIC GENE CODING FOR ITS PRODUCTION
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WHITHAM, CURTIS & WHITHAM
; STREET: Reston Intl. Center, 11800 Sunrise Valley Dr.,
; STREET: Suite 900
```


CITY: Reston
STATE: VA
COUNTRY: USA
ZIP: 20191
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/814,220
FILING DATE: 23-DEC-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/588,437
FILING DATE: 25-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Whitham, Michael E.
REGISTRATION NUMBER: 32,635
REFERENCE/DOCKET NUMBER: CIT.016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-391-2510
TELEFAX: 703-391-9035
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "synthetic DNA"
US-07-814-220-26

Query Match 0.6%; Score 29.4; DB 2; Length 57;
Best Local Similarity 70.9%; Pred. No. 6.6e+02;
Matches 39; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 1468 CGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAGCGGCAGC 1522
Db 1 CCGCAGTTGCTGCCGAGCAGCTGCTGCGGTAGCAGCGGTAGCTGCCGAGCGCGC 55

RESULT 14
US-07-812-421-26
Sequence 26, Application US/07812421
Patent No. 5932697
GENERAL INFORMATION:
APPLICANT: Caceci, Thomas
APPLICANT: Toth, Thomas E.
APPLICANT: Szumanski, Maria B.W.
TITLE OF INVENTION: SYNTHETIC ANTIFREEZE PEPTIDE AND
TITLE OF INVENTION: SYNTHETIC GENE CODING FOR ITS PRODUCTION
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: WHITHAM, CURTIS & WHITHAM
STREET: Reston Intl. Center, 11800 Sunrise Valley Dr.,
STREET: Suite 900
CITY: Reston
STATE: VA
COUNTRY: USA
ZIP: 20191
COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/812,421
FILING DATE: 23-DEC-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/588,437
FILING DATE: 25-SEP-1990

ATTORNEY/AGENT INFORMATION:
NAME: Whitham, Michael E.
REGISTRATION NUMBER: 32,635
REFERENCE/DOCKET NUMBER: CIT.016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-391-2510
TELEFAX: 703-391-9035
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "synthetic DNA"
US-07-812-421-26

Query Match 0.6%; Score 29.4; DB 2; Length 57;
Best Local Similarity 70.9%; Pred. No. 6.6e+02;
Matches 39; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 1468 CGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAGCGGCAGC 1522
Db 1 CCGCAGTTGCTGCCGAGCAGCTGCTGCGGTAGCAGCGGTAGCTGCCGAGCGCGC 55

RESULT 15
US-07-814-220-25/c
Sequence 25, Application US/07814220
Patent No. 5925540
GENERAL INFORMATION:
APPLICANT: Caceci, Thomas
APPLICANT: Toth, Thomas E.
APPLICANT: Szumanski, Maria B.W.
TITLE OF INVENTION: SYNTHETIC ANTIFREEZE PEPTIDE AND
TITLE OF INVENTION: SYNTHETIC GENE CODING FOR ITS PRODUCTION
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: WHITHAM, CURTIS & WHITHAM
STREET: Reston Intl. Center, 11800 Sunrise Valley Dr.,
STREET: Suite 900
CITY: Reston
STATE: VA
COUNTRY: USA
ZIP: 20191
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/814,220
FILING DATE: 23-DEC-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/588,437
FILING DATE: 25-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Whitham, Michael E.
REGISTRATION NUMBER: 32,635
REFERENCE/DOCKET NUMBER: CIT.016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-391-2510
TELEFAX: 703-391-9035
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 62 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "synthetic DNA"
US-07-814-220-25

Fri Apr 29 16:20:38 2005

Query Match 0.6%; Score 29.4; DB 2; Length 62;
Best Local Similarity 70.9%; Pred. No. 6.9e+02;
Matches 39; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 1468 CCGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAGCGGCAGC 1522
Db 62 CCGCAGTGTGCGCAGCAGCTGCTGCGGTAGCAGCGGTAGCTGCCGAGCGGC 8

Search completed: April 29, 2005, 02:23:45
Job time : 762.328 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 07:33:31 ; Search time 2706.88 Seconds
(without alignments)
11433.739 Million cell updates/sec

Title: US-09-674-237B-1
Perfect score: 5084
Sequence: 1 cggcagaggaggaggag.....gagaatcgatcatcaagctt 5084

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5642217 seqs, 3043843248 residues

Total number of hits satisfying chosen parameters: 5511450

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	1.3	65	10	US-09-908-975-27224
2	54	1.1	65	10	US-09-908-975-1543
3	48.8	1.0	60	10	US-09-908-975-10766
4	37.6	0.7	63	10	US-09-854-867-463
5	37.6	0.7	69	19	US-10-786-970A-463
6	36.6	0.7	87	18	US-10-333-894A-6
7	34.6	0.7	63	18	US-10-407-818-8
8	33.4	0.7	93	9	US-09-864-761-24680
9	32.4	0.6	87	19	US-10-759-731A-156
10	32.2	0.6	69	10	US-09-854-867-462
11	32.2	0.6	69	19	US-10-786-970A-462

12	31.2	0.6	77	18	US-10-758-307-66	Sequence 66, Appl
13	31.2	0.6	77	19	US-10-852-797-339	Sequence 339, App
C 14	31.2	0.6	86	14	US-10-085-906-174	Sequence 174, App
15	30.8	0.6	99	16	US-10-029-386-17340	Sequence 17340, A
C 16	30.4	0.6	87	19	US-10-759-731A-155	Sequence 155, App
17	30.4	0.6	96	16	US-10-061-201-11	Sequence 11, Appl
C 18	29.8	0.6	91	16	US-10-029-386-22222	Sequence 22222, A
19	29.8	0.6	81	10	US-09-373-658-52	Sequence 52, Appl
C 20	29.4	0.6	81	11	US-09-989-687-52	Sequence 52, Appl
C 21	29.4	0.6	86	19	US-10-759-731A-158	Sequence 158, App
22	29.4	0.6	97	14	US-10-023-066A-57	Sequence 57, Appl
23	29.4	0.6	97	18	US-10-804-678-57	Sequence 57, Appl
C 24	29.4	0.6	99	18	US-10-021-323-14338	Sequence 14338, A
C 25	29.2	0.6	87	19	US-10-759-731A-157	Sequence 157, App
C 26	29.2	0.6	99	16	US-10-323-051-21	Sequence 21, Appl
C 27	29	0.6	88	19	US-10-488-936-2	Sequence 2, Appl
C 28	29	0.6	96	18	US-10-021-323-14266	Sequence 14266, A
C 29	28.8	0.6	89	18	US-10-021-323-14335	Sequence 14335, A
C 30	28.8	0.6	89	18	US-10-021-323-15706	Sequence 15706, A
31	28.6	0.6	51	18	US-10-865-478-506	Sequence 506, App
32	28.2	0.6	58	16	US-10-275-071-15	Sequence 15, Appl
33	28.2	0.6	84	14	US-10-023-066A-65	Sequence 65, Appl
34	28.2	0.6	84	18	US-10-804-678-65	Sequence 65, Appl
35	28	0.6	68	14	US-10-096-986-58	Sequence 58, Appl
36	28	0.6	69	10	US-09-854-867-468	Sequence 468, App
37	28	0.6	69	19	US-10-786-970A-468	Sequence 468, App
38	28	0.6	75	18	US-10-645-471A-31	Sequence 31, Appl
39	28	0.6	79	17	US-10-260-238-2343	Sequence 2343, Ap
C 40	28	0.6	84	14	US-10-023-066A-66	Sequence 66, Appl
C 41	28	0.6	84	18	US-10-804-678-66	Sequence 66, Appl
C 42	28	0.6	99	17	US-10-072-809A-11	Sequence 11, Appl
C 43	27.8	0.5	88	18	US-10-021-323-808	Sequence 808, App
C 44	27.8	0.5	93	10	US-09-738-937-14	Sequence 14, Appl
45	27.8	0.5	96	17	US-10-353-678-66	Sequence 66, Appl

ALIGNMENTS

RESULT 1
US-09-908-975-27224
; Sequence 27224, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27224
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-27224

Query Match 1.3%; Score 65; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4058 GCAGTTTACCTCATTTGACCTTAGTTGTCATGTCGAATCTCTAGTCACCTGCGCA 4117
DB 1 GCAGTTTACCTCATTTGACCTTAGTTGTCATGTCGAATCTCTAGTCACCTGCGCA 60

QY 4118 GAGGC 4122
 Db 61 GAGGC 65

RESULT 2

US-09-908-975-1543
 ; Sequence 1543, Application US/09908975
 ; Publication No. US20030165843A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SHOSHAN, Avi
 ; APPLICANT: WASSERMAN, Alon
 ; APPLICANT: MINTZ, Eli
 ; APPLICANT: MINTZ, Liat
 ; APPLICANT: FAIGLER, Simchon
 ; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
 ; FILE REFERENCE: 36688-0005
 ; CURRENT APPLICATION NUMBER: US/09/908,975
 ; CURRENT FILING DATE: 2001-07-20
 ; PRIOR APPLICATION NUMBER: US 60/287,724
 ; PRIOR FILING DATE: 2001-05-02
 ; PRIOR APPLICATION NUMBER: US 60/221,607
 ; PRIOR FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 32337
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 1543
 ; LENGTH: 65
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 US-09-908-975-1543

Query Match 1.1%; Score 54; DB 10; Length 65;
 Best Local Similarity 98.5%; Pred. No. 0.00014;
 Matches 65; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 4056 GAGCAGTTACCTTCATTTGACCTTAGTTCATGTCGAAATGTTCTGAGTCACTCGGTG 4115
 Db 1 GAGCAGTTACCTTCATTTGACCTTAGTTCATGTCGAAATGTTCTGAGTCACTCGGTG 59

QY 4116 CAGAGG 4121
 Db 60 CAGAGG 65

RESULT 3

US-09-908-975-10766
 ; Sequence 10766, Application US/09908975
 ; Publication No. US20030165843A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SHOSHAN, Avi
 ; APPLICANT: WASSERMAN, Alon
 ; APPLICANT: MINTZ, Eli
 ; APPLICANT: MINTZ, Liat
 ; APPLICANT: FAIGLER, Simchon
 ; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
 ; FILE REFERENCE: 36688-0005
 ; CURRENT APPLICATION NUMBER: US/09/908,975
 ; CURRENT FILING DATE: 2001-07-20
 ; PRIOR APPLICATION NUMBER: US 60/287,724
 ; PRIOR FILING DATE: 2001-05-02
 ; PRIOR APPLICATION NUMBER: US 60/221,607
 ; PRIOR FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 32337
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 10766
 ; LENGTH: 60
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-908-975-10766

Query Match 1.0%; Score 48.8; DB 10; Length 60;
 Best Local Similarity 88.3%; Pred. No. 0.0039;
 Matches 53; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2489 GAGCGCTGTACCCCTTTGAATCCAGAGTCCAGATGAGATCACCATCCAGCCAGGAGATA 2548
 Db 1 GGGCAGCTGTACCCCTTTGAATCCAGAGCCATGATGAATCACTATCCAGCCAGGAGACA 60

RESULT 4

US-09-854-867-463
 ; Sequence 463, Application US/09854867
 ; Publication No. US20030224356A1
 ; GENERAL INFORMATION:
 ; APPLICANT: JOAN, KNOLL, H
 ; APPLICANT: ROGAN, PETER, K
 ; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
 ; FILE REFERENCE: 30307
 ; CURRENT APPLICATION NUMBER: US/09/854,867
 ; CURRENT FILING DATE: 2003-05-08
 ; NUMBER OF SEQ ID NOS: 613
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 463
 ; LENGTH: 69
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: repetitive sequence found in
 ; OTHER INFORMATION: many eutherian genomes. Length of core repeating element is vari
 ; OTHER INFORMATION: able and is often polymorphic
 US-09-854-867-463

Query Match 0.7%; Score 37.6; DB 10; Length 69;
 Best Local Similarity 72.1%; Pred. No. 6.6;
 Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1455 GCAGGAGGCCAAGCGGCAGCTGAGCTGAGAGAGAGCTGGAGAGAGCGGGAGCTGGA 1514
 Db 1 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 60

QY 1515 GCGGCAGC 1522
 Db 61 GCAGCAGC 68

RESULT 5

US-10-786-970A-463
 ; Sequence 463, Application US/10786970A
 ; Publication No. US2005006449A1
 ; GENERAL INFORMATION:
 ; APPLICANT: JOAN, KNOLL
 ; APPLICANT: ROGAN, PETER
 ; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
 ; FILE REFERENCE: 30307
 ; CURRENT APPLICATION NUMBER: US/10/786,970A
 ; CURRENT FILING DATE: 2004-02-24
 ; PRIOR APPLICATION NUMBER: US/09/573,080
 ; PRIOR FILING DATE: 2000-05-16
 ; NUMBER OF SEQ ID NOS: 479
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 463
 ; LENGTH: 69
 ; TYPE: DNA
 ; ORGANISM: Artificial
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: repetitive sequence found in
 ; OTHER INFORMATION: many eutherian genomes. Length of core repeating element is vari
 ; OTHER INFORMATION: able and is often polymorphic
 US-10-786-970A-463

Query Match 0.7%; Score 37.6; DB 19; Length 69;
 Best Local Similarity 72.1%; Pred. No. 6.6;
 Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
;
;
FEATURE:
OTHER INFORMATION: 3'-amidated
```

```

; OTHER INFORMATION: MAP TO AC021782.2
;
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
;

```

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.9
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.5
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.6
 OTHER INFORMATION: NT HIT: g15031896, EVALUE 2.00e-13
 OTHER INFORMATION: EST_HUMAN HIT: R18580.1, EVALUE 3.00e-13
 US-09-864-761-24680

Query Match 0.7%; Score 33.4; DB 9; Length 93;
 Best Local Similarity 72.9%; Pred. No. 1.3e+02;
 Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1468 CGCAGCTGGAGTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAGCGGAGCGAGGAGA 1526
 Db 93 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGGAAA 35

RESULT 9

US-10-759-731A-156/c
 Sequence 156, Application US/10759731A
 Publication No. US20050079574A1

GENERAL INFORMATION:
 APPLICANT: Bond, Christopher J.

TITLE OF INVENTION: SYNTHETIC ANTIBODY PHAGE LIBRARIES

FILE REFERENCE: 11669.136USU1

CURRENT APPLICATION NUMBER: US/10/759,731A

CURRENT FILING DATE: 2004-01-16

PRIOR APPLICATION NUMBER: US 60/441,059

PRIOR FILING DATE: 2003-01-16

PRIOR APPLICATION NUMBER: US 60/488,610

PRIOR FILING DATE: 2003-07-18

PRIOR APPLICATION NUMBER: US 60/510,314

PRIOR FILING DATE: 2003-10-08

NUMBER OF SEQ ID NOS: 194

SOFTWARE: PatentIn version 3.3

SEQ ID NO 156

LENGTH: 87

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: VLK ala scan

FEATURE:

NAME/KEY: misc feature

LOCATION: (26)..(26)

OTHER INFORMATION: y is c or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (28)..(28)

OTHER INFORMATION: s is g or c

FEATURE:

NAME/KEY: misc feature

LOCATION: (29)..(29)

OTHER INFORMATION: y is c or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (31)..(31)

OTHER INFORMATION: r is a or g

FEATURE:

NAME/KEY: misc feature

LOCATION: (32)..(32)

OTHER INFORMATION: m is a or c

FEATURE:

NAME/KEY: misc feature

LOCATION: (34)..(35)

OTHER INFORMATION: s is g or c

FEATURE:

NAME/KEY: misc feature

LOCATION: (37)..(38)

OTHER INFORMATION: s is g or c

FEATURE:

NAME/KEY: misc feature

LOCATION: (41)..(41)

OTHER INFORMATION: s is g or c

FEATURE:
 NAME/KEY: misc feature
 LOCATION: (43)..(43)
 OTHER INFORMATION: k is g or t
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (46)..(46)
 OTHER INFORMATION: k is g or t
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (50)..(50)
 OTHER INFORMATION: y is c or t
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (53)..(53)
 OTHER INFORMATION: s is g or c
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (56)..(56)
 OTHER INFORMATION: y is c or t
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (58)..(58)
 OTHER INFORMATION: k is g or t
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (59)..(59)
 OTHER INFORMATION: y is c or t
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (61)..(61)
 OTHER INFORMATION: r is a or g
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (64)..(65)
 OTHER INFORMATION: s is g or c
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (68)..(68)
 OTHER INFORMATION: y is c or t
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (70)..(70)
 OTHER INFORMATION: s is g or c
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (71)..(71)
 OTHER INFORMATION: m is a or c
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (73)..(73)
 OTHER INFORMATION: k is g or t
 US-10-759-731A-156

Query Match 0.6%; Score 32.4; DB 19; Length 87;

Best Local Similarity 41.9%; Pred. No. 2.3e+02;

Matches 36; Conservative 22; Mismatches 28; Indels 0; Gaps 0;

Qy 1454 AGCAGGAGGCCAAGCGGAGCTGGAGAGCAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCTGG 1513
 Db 87 ACCCTGACCCCGAGTGTSCARCASSAGYARMARYASCARCGMGCMASCASSASTKYARS 28

Qy 1514 AGCGGAGCGAGAGGAGGAGGAGG 1539

Db 27 ARCAACAGCACCACCAAGTATAGACGG 2

RESULT 10

US-09-854-867-462

	Query Match	0.6%;	Score 31.2;	DB 19;	Length 77;
	Best Local Similarity	70.0%;	Pred. No. 4.7e+02;		
	Matches ' 42;	Conservative 0;	Mismatches 18;	Indels 0;	Gaps 0;
Qy	1444	GAGCGCCAGGAGCAGGAGCGCCAAAGCGCGCAGCTGGAGCTGGAAAGCAGCTGGAGAGACAG	1503		
Db	2	GAGCGCAGAAATCAGGAGATACAGCGGCTCATGGACATCAAGTCGGCGCTGGAGCAGGAG	61		

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	48.8	1.3	60	6	CQ541131	Sequence
C 2	41.8	1.1	98	6	BD270475	Synthetic
C 3	41.8	1.1	98	6	AX039308	Sequence
C 4	41.8	1.1	98	6	AX039517	Sequence
5	41.8	1.1	100	6	BD270476	Synthetic
6	41.8	1.1	100	6	AX039309	Sequence
7	41.8	1.1	100	6	AX039518	Sequence
C 8	40.4	1.1	92	6	BD270477	Synthetic
C 9	40.4	1.1	92	6	AX039310	Sequence
C 10	40.4	1.1	92	6	AX039519	Sequence
11	40.4	1.1	94	6	BD270478	Synthetic
12	40.4	1.1	94	6	AX039311	Sequence
13	40.4	1.1	94	6	AX039520	Sequence
14	38	1.0	78	6	AR159552	Sequence
15	37.8	1.0	82	5	HCH388053	Hyla chry
C 16	37.8	1.0	92	8	AF143273	Vitis vin
17	37	1.0	88	6	BD270480	Synthetic
18	37	1.0	88	6	AX039313	Sequence
19	37	1.0	88	6	AX039522	Sequence

CELLTECH THERAPEUTICS LIMITED (GB)	
FEATURES	Location/Qualifiers
source	1..100
	/organism="synthetic construct"
	/mol_type="unassigned DNA"
	/db_xref="taxon:32630"
	/note="B6464"
ORIGIN	
Query Match	1.1%; Score 41.8; DB 6; Length 100;
Best Local Similarity	68.2%; Pred. No. 2e+02;
Matches	58; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
Qy	1186 GAGCGCCAGGAGGAGGCGCAAGCGGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAG 1245
Db	15 GAGCAGCAGGAGCAGGAGGAGCAGTAGCAGGAGCAGCAGGAGCAGCAGGAGCAGGAGCAGGAG 74
Qy	1246 CGGGAGCTGGAGCGGCGGAGGAGG 1270
Db	75 CAGCAGGAGCAGCAGCAGCCAAAAG 99
RESULT 8	
BD270477/c	
LOCUS	BD270477 92 bp DNA linear PAT 17-JUL-2003
DEFINITION	Synthetic transmembrane components.
ACCESSION	BD270477
VERSION	BD270477.1 GI:33080245
KEYWORDS	JP 2002541845-A/11.
SOURCE	synthetic construct
ORGANISM	other sequences; artificial sequences.
REFERENCE	1 (bases 1 to 92)
AUTHORS	Lawson,A.D.G. and Finney,H.M.
TITLE	Synthetic transmembrane components
JOURNAL	Patent: JP 2002541845-A 11 10-DEC-2002;
COMMENT	CELLTECH THERAPEUTICS LTD
	OS Artificial Sequence
	PN JP 2002541845-A/11
	PD 10-DEC-2002
	PF 17-APR-2000 JP 2000612453
	PR 16-APR-1999 GB 9908816.3,16-APR-1999 GB 9908818.9 PI
	ALASTAIR DAVID GRIFFITHS LAWSON,HELENE MARGARET FINNEY PC
	C12N15/09,A61K38/00,A61K48/00,A61P1/00,A61P3/10,A61P7/06, PC
	A61P11/06,
	PC A61P17/00,A61P17/06,A61P19/02,A61P25/00,A61P29/00,A61P31/18,
	PC A61P35/00,
	PC A61P37/06,A61P37/08,C07K14/705,C12N1/15,C12N1/19,C12N1/21, PC
	C12N5/10,
	PC C12N15/00,C12N5/00,A61K37/02
	CC B6465
	FT Key
	FT source
FT	1..92
	/organism='Artificial Sequence'.
FEATURES	Location/Qualifiers
source	1..92
	/organism="synthetic construct"
	/mol_type="genomic DNA"
	/db_xref="taxon:32630"
ORIGIN	
Query Match	1.1%; Score 40.4; DB 6; Length 92;
Best Local Similarity	71.6%; Pred. No. 4.2e+02;
Matches	53; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
Qy	1179 AGAGCGGAGCGCCAGGAGCAGGAGGCGCAAGCGGAGCTGGAGCTGGAGAGCAGCTGGA 1238
Db	86 AGAACAGGAGCAGCAGGAGCAGTAGCAGGAGCAGCAGCAGGAGCAGCAGGAGCAGGAGCA 27
Qy	1239 GAGCAGCGGGAGC 1252
Db	26 GCAGGAGCAGCAGC 13

RESULT 9					
AX039310/c			92 bp	DNA	linear PAT 18-NOV-2000
LOCUS					
DEFINITION	Sequence 11 from Patent WO0063373.				
ACCESSION	AX039310				
VERSION	AX039310.1 GI:11229415				
KEYWORDS	synthetic construct				
SOURCE	synthetic construct				
ORGANISM	other sequences; artificial sequences.				
1					
REFERENCE	Finney,H.M. and Lawson,A.D.				
AUTHORS					
TITLE	A method of altering the properties of a membrane-associated prote				
	in by substitution of the transmembrane domain				
JOURNAL	Patent: WO 0063373-A 11 26-OCT-2000;				
	CELLTECH THERAPEUTICS LIMITED (GB)				
FEATURES	Location/Qualifiers				
source	1..92				
	/organism="synthetic construct"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32630"				
	/note="B6465"				
ORIGIN					
Query Match	1.1%; Score 40.4; DB 6; Length 92;				
Best Local Similarity	71.6%; Pred.No.4.2e+02;				
Matches	53; Conservative 0; Mismatches 21; Indels 0; Gaps 0;				
QY	1179 AGAGCGGGAGCGCCAGGAGGAGGCCAAGCGGCAGCTGGAGCTGGGAAGCAGCTTGA	1238			
Db	86 AGAACAGGAGCAGCAGCAGCAGTAGCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGGAGCA	27			
QY	1239 GAAGCAGCGGGAGC	1252			
Db	26 GCAGGAGCAGCAGC	13			
RESULT 10					
AX039519/c			92 bp	DNA	linear PAT 18-NOV-2000
LOCUS					
DEFINITION	Sequence 11 from Patent WO0063374.				
ACCESSION	AX039519				
VERSION	AX039519.1 GI:11229546				
KEYWORDS	synthetic construct				
SOURCE	synthetic construct				
ORGANISM	other sequences; artificial sequences.				
1					
REFERENCE	Lawson,A.D. and Finney,H.M.				
AUTHORS					
TITLE	Synthetic transmembrane components				
JOURNAL	Patent: WO 0063374-A 11 26-OCT-2000;				
	CELLTECH THERAPEUTICS LIMITED (GB)				
FEATURES	Location/Qualifiers				
source	1..92				
	/organism="synthetic construct"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32630"				
	/note="B6465"				
ORIGIN					
Query Match	1.1%; Score 40.4; DB 6; Length 92;				
Best Local Similarity	71.6%; Pred.No.4.2e+02;				
Matches	53; Conservative 0; Mismatches 21; Indels 0; Gaps 0;				
QY	1179 AGAGCGGGAGCGCCAGGAGGAGGCCAAGCGGCAGCTGGAGCTGGGAAGCAGCTTGA	1238			
Db	86 AGAACAGGAGCAGCAGCAGCAGTAGCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGGAGCA	27			
QY	1239 GAAGCAGCGGGAGC	1252			
Db	26 GCAGGAGCAGCAGC	13			

Search completed: April 28, 2005, 18:48:48
Job time : 15158.7 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 05:12:48 ; Search time 1763.82 Seconds
(without alignments)

12223.299 Million cell updates/sec

Title: US-09-674-237B-2

Perfect score: 3642

Sequence: 1 atggctcagtttccacacc.....tggaccgccagcgaatga 3642

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 4530610

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	2.2	80	12 ADP49483	Adp49483 Oligonuc1
2	48.8	1.3	60	6 ABN38018	Abn38018 Human spl
3	42.8	1.2	81	10 ADC16446	Adc16446 Short int
4	42.8	1.2	81	10 ADC16445	Adc16445 Short int
5	41.8	1.1	98	4 AAC65404	Aac65404 Oligonuc1
6	41.8	1.1	100	4 AAC65405	Aac65405 Oligonuc1
7	41	1.1	75	6 ABK89299	Abk89299 Trinucleo
8	40.4	1.1	92	4 AAC65406	Aac65406 Oligonuc1
9	40.4	1.1	94	4 AAC65407	Aac65407 Oligonuc1
10	39.2	1.1	97	10 ACD94308	Acd94308 Human col
11	39	1.1	89	2 AAV68382	Av68382 Clone #8
12	38	1.0	78	2 AAV17232	Av17232 SCA2 gene
13	38	1.0	78	2 AAV17233	Av17233 SCA2 gene
14	37.6	1.0	69	7 ADS31430	Ads31430 Human gen
15	37	1.0	88	4 AAC65409	Aac65409 Oligonuc1
16	36.6	1.0	87	6 AAD30431	Aad30431 29 mer ol
17	36.4	1.0	78	2 AAV17231	Av17231 SCA2 gene
18	35.8	1.0	86	4 AAC65408	Aac65408 Oligonuc1
19	34	0.9	75	2 AAV17230	Av17230 SCA2 gene
20	33.4	0.9	93	4 AAI26445	Aai26445 Probe #16

C 21	33.4	0.9	93	4	ABA74544	Human foe
C 22	33.4	0.9	93	4	AAI55023	Probe #23
C 23	33.4	0.9	93	4	ABA39360	Probe #17
C 24	33.4	0.9	93	4	AAK49188	Human bon
C 25	33.4	0.9	93	4	AAK23011	Human bra
C 26	33.4	0.9	93	4	ABS48832	Human liv
C 27	33.4	0.9	93	6	ABS22756	Human gen
C 28	33	0.9	66	2	AAW78908	Poly-glut
C 29	33	0.9	69	2	AAV17234	SCA2 gene
C 30	32.4	0.9	87	13	ADQ95148	Adq95148 Synthetic
C 31	32.2	0.9	69	7	ADS31429	Human gen
C 32	31.4	0.9	77	4	AAK65410	Oligonuc1
C 33	31.4	0.9	79	4	AAK65411	Oligonuc1
C 34	31.2	0.9	77	13	ADR00028	Adr00028 KRT19 PCR
C 35	31	0.9	75	12	ADH58952	Silk prot
C 36	30.8	0.8	99	12	ACH84145	Human gen
C 37	30.6	0.8	78	2	AAV27572	Nucleotid
C 38	30.4	0.8	87	13	ADQ95147	Synthetic
C 39	30.4	0.8	96	6	ABV89298	Human POS
C 40	30.4	0.8	98	4	AAI84741	Human pol
C 41	30	0.8	86	3	AAC10510	Human sec
C 42	29.8	0.8	91	12	ACH89027	Human gen
C 43	29.4	0.8	57	2	AAK88107	Antifreez
C 44	29.4	0.8	62	2	AAK88106	Synthetic
C 45	29.4	0.8	75	12	ADH58951	Silk prot

ALIGNMENTS

RESULT 1

ADP49483

ID ADP49483 standard; DNA; 80 BP.

XX ADP49483;

AC ADP49483;

XX DT 12-AUG-2004 (first entry)

XX DE Oligonucleotide array related rat oligonucleotide probe No 99.

XX KW oligonucleotide array; orthologue; homology; expression distribution;

XX OS Rattus norvegicus.

XX PN JP2004016070-A.

XX PD 22-JAN-2004.

XX PF 14-JUN-2002; 2002JP-00174208.

XX PR 14-JUN-2002; 2002JP-00174208.

XX PA (HITA) HITACHI LTD.

XX DR WPI; 2004-113862/12.

XX PT Oligonucleotide array, useful for measuring ortholog gene-expression

XX PT distribution, comprising number of oligonucleotides derived from ortholog

XX PT gene of different organism immobilized on support.

XX PS Disclosure; Page 16; 56pp; Japanese.

XX CC The invention relates to a novel oligonucleotide array comprising a

XX CC number of oligonucleotides derived from an orthologue gene of a different

XX CC organism or species immobilised on a support body. The oligonucleotide

XX CC array has two sides comprising a gene derived from two different

XX CC organisms, in which one side comprises a human gene. In the

XX CC oligonucleotide array a base sequence differs in a different position on

XX CC the support body. Each of the oligonucleotides is a partial sequence of

XX CC the orthologue gene of the organism A and organism B. The sequence match

XX CC degree of the organism species A and the organism species B is less than

XX CC 70%. The sequence homology of the genes other than the orthologue gene of

CC the organism species A and the organism species B is the level-of-
CC statistical-significance value, calculated using the homology searching
CC algorithm: basic local alignment search tool (BLAST) and the value is 0.1
CC or more. The found value or the calculated value of the melting-
CC temperature of the variant genes other than the orthologue gene of the
CC organism species A and the organism species B is 20 degrees C or more.
CC The oligonucleotide array is useful for measuring expression distribution
CC of the orthologue gene in different organisms. The method is useful for
CC comparing expression change of the orthologue gene with respect to
CC medical-agent administration and for comparing change of the function of
CC the orthologue gene in different organisms. The oligonucleotide array has
CC the ability to perform a measurement of gene-expression distribution of
CC two or more types of organism simultaneously. The oligonucleotide array
CC improves reliability of measurement. This polynucleotide sequence
CC represents a probe of a rat gene for comparison against a human gene used
CC in the oligonucleotide array of the invention.

XX Sequence 80 BP; 19 A; 15 C; 25 G; 21 T; 0 U; 0 Other;

Query Match 2.2%; Score 80; DB 12; Length 80;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2817 AGACATGTGTGTTGGAGAGTTCAAGTTCAGAGGGTGTGTTCCCAAGTCTTACGT 2876
Db 1 AGACATGTGTGTTGGAGAGTTCAAGTTCAGAGGGTGTGTTCCCAAGTCTTACGT 60

QY 2877 GAAACTCATTTACGGCCCG 2896
Db 61 GAAACTCATTTACGGCCCG 80

RESULT 2
ID ABN38018
XX ABN38018 standard; DNA; 60 BP.
AC ABN38018;
XX
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:10766.
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Homo sapiens.
XX
XX WO200210449-A2.
XX
XX 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-IB001903.
XX
XX 28-JUL-2000; 2000US-0221607P.
XX
XX 02-MAY-2001; 2001US-0287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of a
PT genome, useful for detecting tissue-, pathology-, and developmental-
PT specific genes.
XX
XX Example 1; SEQ ID NO 10766; 47pp; English.

PS The present invention describes oligonucleotide libraries for detecting

XX messenger RNAs that populate a (sub-)transcriptome, where the (sub-
CC)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises several

CC oligonucleotides, each capable of hybridising selectively to a set of
CC messenger RNAs transcribed from a given transcription unit of the genome,
CC which encodes one or more messenger RNA splice variants. The
CC oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a particular
CC biological or pathological state, and so allowing the detection of tissue
CC - and pathology-specific genes such as those genes only expressed in
CC specific tissue under a specific pathological condition; to detect
CC developmental specific genes; and to detect RNA transcripts and splice
CC variants of a transcriptome of a patient suffering from a particular
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC rats, humans and mice, which are used in the exemplification of the
CC present invention. N.B. The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 60 BP; 19 A; 17 C; 13 G; 11 T; 0 U; 0 Other;

Query Match 1.3%; Score 48.8; DB 6; Length 60;
Best Local Similarity 88.3%; Pred. No. 0.13;
Matches 53; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2231 GAGCGTGATCCCTTTGAATCCAGAGTCACGATGAGATCACCATCCAGCCAGGAGATA 2290
Db 1 GGGCACTGTACCCCTTTGAATCCAGAGTCACGATGAGATCACCATCCAGCCAGGAGATA 60

RESULT 3
ID ADC16446/C
XX ADC16446 standard; RNA; 81 BP.
AC ADC16446;
XX
XX 18-DEC-2003 (first entry)
XX
DE Short interfering double-stranded RNA oligonucleotide SEQ ID NO:171.
XX
XX expression interference; expression inhibition; target gene;
KW short interfering double stranded RNA; cytostatic; gene therapy;
KW proliferative disease; cancer; ds.
XX
XX Synthetic.
XX
XX WO2003012052-A2.
XX
XX 13-FEB-2003.
XX
XX 30-JUL-2002; 2002WO-US024226.
XX
XX 30-JUL-2001; 2001US-0308640P.
XX
XX 08-APR-2002; 2002US-0370970P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (CARN-) CARNegie INST WASHINGTON.
XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
XX Caplen NJ, Morgan RA, Fire A, Parrish S, Mouses S;
XX Kallioniemi O, Cornelison JR, Alton EW, Griesenbach U;
XX WPI; 2003-248169/24.
XX
XX New RNA comprising double stranded RNA and a 3' or 5' overhang having a
PT length of 0-nucleotide to 5-nucleotides on each strand, useful as reverse
PT genetic and/or therapeutic tools for interfering or inhibiting expression
PT of a target gene.

PS Claim 71; SEQ ID NO 171; 176pp; English.

XX The present invention describes an RNA (I) used for the interference or

PT mammalian cells useful for identifying disorders associated with TNR
 PT instability such as cancer, Fragile X syndrome, Huntington's disease or
 XX myotonic dystrophy.
 XX
 XX Example 1; Page 24; 43pp; English.
 XX
 CC The invention relates to detecting (M1) alterations in trinucleotide
 CC repeat (TNR) tract lengths comprising: (i) contacting mammalian cells
 CC with a shuttle vector (containing approximately 25 repeats) that enters
 CC and replicates in the cells; (ii) recovering the replicated shuttle
 CC vector from the cells; (iii) introducing the vector into a yeast cell in
 CC the presence of a selection agent or (where the expanded TNR confers a
 CC His⁺ phenotype) in the absence of histidine, alteration to the TNR tract
 CC confers resistance to the selective agent or His⁺ phenotype; and (iv)
 CC selecting yeast cells with the shuttle vector containing TNR tract
 CC alterations that survived the presence of the selective agent. Also
 CC included is an adaptation of the used to detect the contraction of 33 or
 CC 55 repeats using growth of the yeast cells in the absence of uracil. The
 CC methods of the present invention are useful for identifying disorders
 CC with genetic alterations associated with TNR instability such as cancer
 CC (e.g. testicular and prostate), Fragile X syndrome, Huntington's disease,
 CC myotonic dystrophy, spinal and bulbar muscular atrophy (SBMA),
 CC spinocerebellar ataxias (types 1, 8, 12, 3 (Machado-Joseph disease), 6,
 CC 7 and 2), dentatorubral-pallidoluysian atrophy (DRPLA), Friedreich's
 CC ataxia and Kennedy's disease. The present sequence is the CAG25 TNR known
 CC to be unstable in Human sperm and in yeast which was tested in the method
 CC of the invention for expansion
 XX
 SQ Sequence 75 BP; 25 A; 25 C; 25 G; 0 T; 0 U; 0 Other;

Query Match 1.1%; Score 41; DB 6; Length 75;
 Best Local Similarity 72.6%; Pred. No. 10;
 Matches 53; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
 QY 1192 CAGCAGCAGGCGCCAGCGCGAGCTGGAGTGGAGCAGCTGGAGAGCAGCGGAG 1251
 DB 1 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 60
 QY 1252 CTGGAGCGCGAGC 1264
 DB 61 CAGCAGCAGCAGC 73

RESULT 8
 AAC65406/C
 ID AAC65406 standard; DNA; 92 BP.
 XX
 AC AAC65406;
 XX
 DT 14-FEB-2001 (first entry)
 XX
 DE Oligonucleotide B6465 for chimeric receptor construction.
 XX
 KW Membrane-associated protein; antiviral; antibacterial; antiparasitic;
 KW immunomodulatory; anticancer; antiinflammatory; antiasthmatic;
 KW antidiabetic; neuroprotective; chimeric receptor; infection;
 KW inflammatory disorder; autoimmune disorder; cancer; allergy; asthma;
 KW eczema; cystic fibrosis; sickle cell anaemia; psoriasis;
 KW multiple sclerosis; organ transplant rejection; diabetes;
 KW transmembrane domain; ss.

Synthetic.
 OS
 XX
 XX WO200063373-A1.
 XX
 XX 26-OCT-2000.
 XX
 XX 17-APR-2000; 2000WO-GB001471.
 XX
 XX 16-APR-1999; 99GB-00008816.
 XX
 XX (CLLT) CELLTech THERAPEUTICS LTD.
 XX
 XX

PI Finney HM, Lawson ADG;
 XX
 DR WPI; 2001-015774/02.
 XX
 XX Altering the properties or level of expression of membrane-associated
 PT proteins, e.g., to change responses to cell surface antigens or the
 PT sensitivity of intracellular signaling.
 XX
 PS Example 2; Fig 3; 47pp; English.
 XX
 CC The present sequence was used in the construction of chimeric receptors.
 CC The properties and level of expression of a membrane-associated protein
 CC may be altered by substituting a transmembrane region or a membrane-
 CC anchoring region for transmembrane or membrane-anchoring regions that are
 CC not naturally part of the protein. The relative response of membrane-
 CC associated proteins to cell surface-associated antigen versus antigen in
 CC solution, and the sensitivity of intracellular signaling mediated by
 CC membrane-associated proteins can be altered. They can be important in
 CC treatment of, e.g. HIV infection, bacterial infections, parasitic
 CC infections, inflammatory/autoimmune disorders (e.g. rheumatoid arthritis,
 CC osteoarthritis or inflammatory bowel disease), cancer, allergic/atopic
 CC diseases (e.g. asthma or eczema), congenital disorders (e.g. cystic
 CC fibrosis or sickle cell anaemia), dermatological disorders (e.g.
 CC psoriasis), neurological disorders (e.g. multiple sclerosis), organ
 CC transplant rejection, graft-versus-host diseases, or metabolic/idiopathic
 CC diseases (e.g. diabetes)
 XX
 SQ Sequence 92 BP; 2 A; 34 C; 24 G; 32 T; 0 U; 0 Other;

Query Match 1.1%; Score 40.4; DB 4; Length 92;
 Best Local Similarity 71.8%; Pred. No. 16;
 Matches 53; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
 QY 1179 AGAGCGGAGCGCCAGCAGCAGCAGCAGCGCGAGCTGGAGTGGAGAGCAGCTGGA 1238
 DB 86 AGAACAGAGCAGCAGCAGCAGCTAGCAGGAGCAGCAGCAGCAGCAGCAGCAGCA 27
 QY 1239 GAAGCAGCGCGAGC 1252
 DB 26 GCAGGAGCAGCAGC 13

RESULT 9
 AAC65407
 ID AAC65407 standard; DNA; 94 BP.
 XX
 AC AAC65407;
 XX
 DT 14-FEB-2001 (first entry)
 XX
 DE Oligonucleotide B6466 for chimeric receptor construction.
 XX
 KW Membrane-associated protein; antiviral; antibacterial; antiparasitic;
 KW immunomodulatory; anticancer; antiinflammatory; antiasthmatic;
 KW antidiabetic; neuroprotective; chimeric receptor; infection;
 KW inflammatory disorder; autoimmune disorder; cancer; allergy; asthma;
 KW eczema; cystic fibrosis; sickle cell anaemia; psoriasis;
 KW multiple sclerosis; organ transplant rejection; diabetes;
 KW transmembrane domain; ss.

Synthetic.
 OS
 XX
 XX WO200063373-A1.
 XX
 XX 26-OCT-2000.
 XX
 XX 17-APR-2000; 2000WO-GB001471.
 XX
 XX 16-APR-1999; 99GB-00008816.
 XX
 XX (CLLT) CELLTech THERAPEUTICS LTD.
 XX
 XX Finney HM, Lawson ADG;

OS Homo sapiens.
XX XX
XX WO9849345-A1.
XX XX
XX PD 05-NOV-1998.
XX XX
XX PF 29-APR-1998; 98WO-US008616.
XX XX
XX PR 29-APR-1997; 97US-0045078P.
XX XX
XX PA (UYBO-) UNIV BOSTON.
XX XX
XX PI Smith CL;
XX XX
XX DR WPI; 1998-594983/50.
XX XX
XX PT Analysing nucleic acid samples - using amplification primers which
PT contain CAG or CTG tri-nucleotide repeats for differential display of
PT samples from different sources.
XX XX
XX Example; Page 32; 44pp; English.
XX XX
XX CC This sequence represents a fragment of a human CAG repeat containing
CC clone DNA sequence isolated using the method of the invention. The method
CC is for analysing nucleic acids in a sample, and comprises: (a) providing
CC a sample containing nucleic acid, a first oligonucleotide primer
CC comprising a CTG repeat, a second oligonucleotide primer comprising a CAG
CC repeat and a polymerase and PCR reagents; (b) preparing the nucleic acid
CC so that it is amplifiable; (c) amplifying the nucleic acid with the first
CC and second primers; and (d) detecting the amplified product. The method
CC is used to distinguish between the expression of genes in two or more
CC biological samples, e.g. body fluids, cells, solid tissue or solid and
CC liquid foods. It can be used in medical diagnostics, e.g. to
CC differentiate between normal and diseased tissue or to assess the
CC variation within monozygotic twin pairs. The method allows the isolation
CC and analysis of genome subsets containing CAG repeats which are known to
CC be important in a number of neurological diseases including Huntington's
CC chorea. The method uses PCR suppression, in which only fragments which
CC contain a target repeat are efficiently amplified. This allows accurate
CC identification of differentially expressed genes in various cell types.
CC Genome complexity is reduced by the new method which targets genomic
CC subsets containing CAG repeats

XX XX
SQ Sequence 89 BP; 2 A; 23 C; 37 G; 27 T; 0 U; 0 Other;
Query Match 1.1%; Score 39; DB 2; Length 89;
Best Local Similarity 65.5%; Pred.No. 33;
Matches 57; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1183 CGGGAGCGCAGGACGAGCAGGAGCGGCCAAGCGGAGCTGGAGAGCAGCTGGAGAAG 1242
DB 89 CAGCAGCAGCAGCAGCAGCAGCAGCCCCACCACGAGCAGCAGCAGCAGCAGCAGCAG 30

QY 1243 CAGCGGAGCTGGAGCGGAGCGGAGCAGGAG 1269
DB 29 CAGCCCCAGCAGCAGCAGCAGCAGCAGCAG 3

RESULT 12
AAVI7232
ID ID AAVI7232 standard; DNA; 78 BP.
XX XX
XX AC AAVI7232;
XX XX
DT 29-JUN-1998 (first entry)
XX XX
DE SCA2 gene CAG repeat unit fragment.
XX XX
KW SCA2 gene; spinocerebellar ataxis type II; CAG repeat; PCR primer; ss.
XX XX
OS Synthetic.

XX XX
PN WO9803679-A1.
XX XX
PD 29-JAN-1998.
XX XX
PF 18-JUL-1996; 96WO-JP001999.
XX XX
PR 18-JUL-1996; 96WO-JP001999.
XX XX
PA (SRLS-) SRL INC.
XX XX
PI Tsuji S, Sanpei K;
XX XX
DR WPI; 1998-120796/11.
XX XX
PT Diagnosing spinocerebellar ataxis type II; CAG repeat; PCR primer; ss.
XX XX
PS Synthetic.

XX XX
PN WO9803679-A1.
XX XX
PP Disclosure; Page 13-14; 23pp; Japanese.

XX XX
PD 29-JAN-1998.
XX XX
XX 18-JUL-1996; 96WO-JP001999.
XX XX
XX 18-JUL-1996; 96WO-JP001999.
XX XX
XX (SRLS-) SRL INC.
XX XX
XX Tsuji S, Sanpei K;
XX XX
XX WPI; 1998-120796/11.
XX XX
XX Diagnosing spinocerebellar ataxis type II - by PCR and determining number
XX of CAG repeat units.
XX XX
XX Disclosure; Page 13; 23pp; Japanese.
XX XX
XX This sequence represents a fragment of the SCA2 gene. It can be used in
XX the method of the invention for diagnosing spinocerebellar ataxis type
XX II, by performing PCR on the test DNA using two primers hybridising to
XX parts of the SCA2 gene sequence, and determining the number of CAG
XX repeats in the amplified products. The method provides an easy means for
XX the diagnosis of spinocerebellar ataxis type II
XX SQ Sequence 78 BP; 23 A; 30 C; 25 G; 0 T; 0 U; 0 Other;
Query Match 1.0%; Score 38; DB 2; Length 78;
Best Local Similarity 67.9%; Pred.No. 55;
Matches 53; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1192 CAGGAGCAGGAGCGCCCAAGCGGAGCTGGAGAGCAGCTGGAGAAGCAGCAGCGGAG 1251
DB 1 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 60

QY 1252 CTGGAGCGGAGCAGCGAGAG 1269
DB 61 CAGCAGCGCGCGCGCGCG 78

RESULT 13
AAVI7233
ID ID AAVI7233 standard; DNA; 78 BP.
XX XX
XX AC AAVI7233;
XX XX
DT 29-JUN-1998 (first entry)
XX XX
DE SCA2 gene CAG repeat unit fragment.
XX XX
KW SCA2 gene; spinocerebellar ataxis type II; CAG repeat; PCR primer; ss.
XX XX
OS Synthetic.

XX XX
PN WO9803679-A1.
XX XX
PD 29-JAN-1998.
XX XX
PF 18-JUL-1996; 96WO-JP001999.
XX XX
PR 18-JUL-1996; 96WO-JP001999.
XX XX
PA (SRLS-) SRL INC.
XX XX
PI Tsuji S, Sanpei K;
XX XX
DR WPI; 1998-120796/11.
XX XX
PT Diagnosing spinocerebellar ataxis type II - by PCR and determining number
XX of CAG repeat units.
XX XX
XX Disclosure; Page 13-14; 23pp; Japanese.

will hybridize to a reduced sample copy. The single copy sequence is deduced by comparing the target nucleic acid (e.g. a disease causing gene) with a collection of high and low complexity repeat sequences as found in the genome of the organism from containing the target nucleic acid. The probe is generated by PCR on the target sequence. The probe is essentially free of blocking nucleic acid sequences which will hybridize to repeat sequences within the genome of which the TNA is a part, and is labelled with a label selected from fluorochrome-responsive labels, fluorochromes, calorimetric chemical, conjugated proteins, antibodies, antigens and

Search completed: April 28, 2005, 08:43:21
Job time : 1764.82 secs

THIS PAGE BLANK (USPTO)

[illegible]

```

|||||
61 AACGCCACGACAAACTATGAGTGACACTTAA 94

RESULT 2
AA589590
LOCUS
DEFINITION
H.sapiens DNA for trapped exon (ID HMC18D03), genomic survey
sequence.
64 bp DNA linear GSS 29-MAY-1997
ACCESSION
X88324
VERSION
X88324.1 GI:1437729
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 64)
AUTHORS
Chen,H., Christ,R., Rossier,C., Morris,M.A., Lalioti,M.D. and
Antonarakis,S.E.
TITLE
Cloning of 559 potential exons of genes of human chromosome 21 by
exon trapping
JOURNAL
Genome Res. 6 (8), 747-760 (1996)
MEDLINE
97011340
PUBMED
8858350
REFERENCE
2 (bases 1 to 64)
AUTHORS
Chen,H.M., Rossier,C., Christ,R. and Antonarakis,S.E.
TITLE
Cloning of trapped exons from human chromosome 21
JOURNAL
Unpublished
REFERENCE
3 (bases 1 to 64)
AUTHORS
Antonarakis,S.E.
TITLE
Direct Submission
JOURNAL
Submitted (17-MAR-1995) Stylianos E. Antonarakis, Division of
Medical Genetics, University and Cantonal Hospital of Geneva, CMU,
1 rue Michel-Servet, 1211 Geneva, SWITZERLAND

FEATURES
source
1..64
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
1..64
/feature="trapped exon"

ORIGIN
Query Match 1.5%; Score 54.6; DB 9; Length 64;
Best Local Similarity 89.1%; Pred. No. 0.027; 7; Indels 0; Gaps 0;
Matches 57; Conservative 0; Mismatches 0;

QY 725 GTCCCCAGGCAAGAACTATTCTCATGCAATCAAGTTTACCCAGGCTCAGCTGGTTCAA 784
Db 1 GTCCCAAGCAAGAACTATTCTCATGCAATCAAGTTTACCCAGGCTCAGTGGTTCAA 60

QY 785 TATG 788
Db 61 TATG 64

RESULT 4
CC200157
LOCUS
DEFINITION
XG352 BayGenomics Gene Trap Library pGTL1xf Mus musculus cDNA, mRNA
sequence.
46 bp mRNA linear GSS 09-MAY-2003
ACCESSION
CC200157
VERSION
CC200157.1 GI:30479920
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 46)
AUTHORS
BayGenomics.
TITLE
http://baygenomics.ucsf.edu/
JOURNAL
Unpublished (2001)
COMMENT
Contact: BayGenomics
Bay Area Functional Genomics Consortium (BayGenomics)
Email: info@baygenomics.ucsf.edu
Sequence tag generated by 5' RACE of total RNA from gene trap ES
cell line. ES cell lines harboring insertion mutation of target
gene are available upon request from BayGenomics. Annotation

|||||
61 AACGCCACGACAAACTATGAGTGACACTTAA 94

AA589590
79 bp mRNA linear EST 16-SEP-1997
v149c09.s1 Stragatene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:975568 3', similar to SW:VFJ4 YEAST P43603 HYPOTHETICAL 40.4
KD PROTEIN IN PES4-H152 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION
AA589590
VERSION
AA589590.1 GI:2402970
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 79)
AUTHORS
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,F., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE
The WashU-HMI Mouse EST Project
JOURNAL
Unpublished (1996)
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:556296
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 1.
FEATURES
source
1..79
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:975568"
/sex="females"
/tissue_type="whole skin"
/dev_stage="11 weeks old"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse skin (#937313)"
/note="Organ: skin; Vector: pBluescript SK-; Site: 1:
ECORI; Site2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb; Uni-ZAP XR vector; -5'
adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor
sequence: 5' CTCGAGTGTGTTTTTTTTTTTTTTT 3'"
1..79
/feature="trapped exon"

ORIGIN
Query Match 2.0%; Score 74.2; DB 1; Length 79;
Best Local Similarity 96.2%; Pred. No. 3.8e-07;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3003 TGCCATGTACATACGAGAGTCTGAGCAGGAGATTAACTTTCAGCAAGGGGATGT 3062
Db 1 TGCCATGTACATACGAGAGTCTGATCAAGGAGATTAAAGTTTCAGCAAGGGGATGT 60

QY 3063 GATTGTGTTTACCAAGAAA 3081
Db 61 GATTGTGTTTACCAAGAAA 79

RESULT 3
HSMC18D03

```


Query Match	Score	DB 1: Length 86;	1.1%:
	40.6:	DB 1:	Length 86;

'RescueMu.' Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 1.0%; Score 36; DB 8; Length 91;
Best Local Similarity 64.3%; Pred. No. 1.2e+03;
Matches 54; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1116 GGAGCAGCAGCGCAAGAGCAGGAGCGTTGGCTCAGCTGGAGCGCGCCGAGCAGAG 1175
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 GGAGGAGACCGCGCAAGAGCAGGAGCGTTAAAGAGGAGGAGGAGGTGGAGCGCGGGGAGAG 26
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1176 GAAAGAGCGGAGCGCCGAGGAGCA 1199
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 GAAGACCCAGGAGCGCGCCGAGCA 2

RESULT 13

BH225515/c
LOCUS BH225515 91 bp DNA linear GSS 08-NOV-2001
DEFINITION 1006126G04.x1 1006 - RescueMu Grid G Zea mays genomic, genomic
survey sequence.

ACCESSION BH225515
VERSION BH225515.1 GI:16823551
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 91)

REFERENCE

AUTHORS Walbot, V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site so sequence was trimmed. Post-ligation
sequence submitted separately.
Plate: 1006126 row: 1
Class: transposon-tagged.

FEATURES

source
Location/Qualifiers
1..91
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1006 - RescueMu Grid G"
/note="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for

'RescueMu.' Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 1.0%; Score 36; DB 8; Length 91;

Best Local Similarity 64.3%; Pred. No. 1.2e+03;
Matches 54; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1116 GGAGCAGCAGCGCAAGAGCAGGAGCGTTGGCTCAGCTGGAGCGCGCCGAGCAGAG 1175
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 GGAGGAGACCGCGCAAGAGCAGGAGCGTTAAAGAGGAGGAGGAGGTGGAGCGCGGGGAGAG 26
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1176 GAAAGAGCGGAGCGCCGAGGAGCA 1199
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 GAAGACCCAGGAGCGCGCCGAGCA 2

RESULT 14

BH225549/c
LOCUS BH225549 91 bp DNA linear GSS 08-NOV-2001
DEFINITION 1006126H07.x1 1006 - RescueMu Grid G Zea mays genomic, genomic
survey sequence.

ACCESSION BH225549
VERSION BH225549.1 GI:16823620
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 91)

REFERENCE

AUTHORS Walbot, V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site so sequence was trimmed. Post-ligation
sequence submitted separately.
Plate: 1006126 row: 1
Class: transposon-tagged.

FEATURES

source
Location/Qualifiers
1..91
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1006 - RescueMu Grid G"
/note="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for

'RescueMu.' Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 1.0%; Score 36; DB 8; Length 91;
Best Local Similarity 64.3%; Pred. No. 1.2e+03;
Matches 54; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1116 GGAGCAGCAGCGCAAGAGCAGGAGCGTTGGCTCAGCTGGAGCGCGCCGAGCAGAG 1175
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 GGAGGAGACCGCGCAAGAGCAGGAGCGTTAAAGAGGAGGAGGAGGTGGAGCGCGGGGAGAG 26
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1176 GAAAGAGCGGAGCGCCGAGGAGCA 1199

```

Db      25  GAAGACCCAGGAGCCGCCAGCA 2
      |||  |||||  |||||  |||||
      CR272899      96 bp  DNA  linear  GSS 06-JUL-2004
      Forward strand read from insert in 5'HPRT insertion targeting and
      chromosome engineering clone MHPN79102, genomic survey sequence.
      CR272899
      CR272899.1  GI:50051754
      GSS; genome survey sequence; MICER.
      Mus musculus (house mouse)
      Mus musculus
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
      1 (bases 1 to 96)
      Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
      Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
      Rogers,J. and Bradley,A.
      Direct Submission
      Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
      CB10 1SA, UK. http://www.sanger.ac.uk/MICER
      FEATURES
      source
      1..96
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10090"
      /clone="MHPN79102"
      /clone_lib="MHPN"

      ORIGIN
      Query Match      1.0%; Score 35.6; DB 9; Length 96;
      Best Local Similarity 62.2%; Pred. No. 1.6e+03;
      Matches 56; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
      QY 1207 AACGGCAGCTGGAGTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAGCGGCGAGCGA 1266
      Db 95 AAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 36
      QY 1267 GAGGAGGAGAGAGAGAGAGAGAGATCGAGAGG 1296
      Db 35 AAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6

      Search completed: April 29, 2005, 02:01:58
      Job time : 10843.9 secs

```


Result No.	Query #			DB	ID	Description
	Score	Match	Length			
1	38	1.0	78	3	US-09-043-303-12	Sequence 12, Appl
2	37.6	1.0	69	4	US-09-573-080A-463	Sequence 463, App
3	36.4	1.0	78	3	US-09-043-303-11	Sequence 11, Appl
4	34	0.9	75	3	US-09-043-303-10	Sequence 10, Appl
5	33	0.9	69	3	US-09-043-303-13	Sequence 13, Appl
6	32.2	0.9	69	4	US-09-573-080A-462	Sequence 462, App
C 7	31.4	0.9	89	4	US-09-270-767-4823	Sequence 4823, A
C 8	31.4	0.9	89	4	US-09-270-767-20105	Sequence 20105, A
C 9	30.4	0.8	93	3	US-08-556-978B-24	Sequence 24, Appl
C 10	30.4	0.8	93	3	US-08-556-978B-25	Sequence 25, Appl
C 11	30.2	0.8	51	1	US-08-068-747-1	Sequence 1, Appl
12	30	0.8	86	4	US-09-513-999C-14585	Sequence 14585, A
13	29.4	0.8	57	2	US-07-814-220-26	Sequence 26, Appl
14	29.4	0.8	57	2	US-07-814-421-26	Sequence 26, Appl
C 15	29.4	0.8	62	2	US-07-814-220-25	Sequence 25, Appl
C 16	29.4	0.8	62	2	US-07-812-421-25	Sequence 25, Appl
17	29.4	0.8	97	1	US-08-182-175A-48	Sequence 48, Appl
18	29.4	0.8	97	1	US-08-474-633A-57	Sequence 57, Appl
19	29.4	0.8	97	3	US-08-823-771-57	Sequence 57, Appl
20	29.4	0.8	97	5	PCU-US92-06412-48	Sequence 48, Appl
C 21	29.2	0.8	99	4	US-09-403-533-21	Sequence 21, Appl
C 22	29	0.8	96	3	US-09-281-481A-6	Sequence 6, Appl
C 23	29	0.8	96	3	US-09-281-481A-7	Sequence 7, Appl
24	28.6	0.8	78	4	US-09-513-999C-35527	Sequence 3527, A
25	28.4	0.8	94	4	US-09-513-999C-29269	Sequence 29269, A
26	28.2	0.8	57	3	US-09-043-303-14	Sequence 14, Appl
27	28.2	0.8	58	3	US-08-860-038-15	Sequence 15, Appl


```
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 46; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1218 GGAGCTGGAGAAGCAGCTGGAGAAGCAGCGGAGCTGGAGCGGCGAGGAGGAGAG 1277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1278 GAGGAGGAGGA 1286
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GGAGGAGGAGGA 69

RESULT 7
US-09-270-767-4823/c
; Sequence 4823, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4823
; LENGTH: 89
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-4823

Query Match 0.9%; Score 31.4; DB 4; Length 89;
Best Local Similarity 61.7%; Pred. No. 2.2e+02;
Matches 50; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1209 GCGGCGAGCTGGAGTGGAGAAGCAGCTGGAGAAGCAGCGGAGCTGGAGCGGCGAGGA 1268
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 81 GCGGCGCTGGAGCTGGAGGAGGAGCGGAGATCACCGGCTGCTGGAGGAGGAGGAC 22
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1269 GGAGGAGGAGGAGGAGAT 1289
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 TGCTGAATCGGAGCTGGAGCT 1
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-09-270-767-20105/c
; Sequence 20105, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20105
; LENGTH: 89
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-20105

Query Match 0.9%; Score 31.4; DB 4; Length 89;
Best Local Similarity 61.7%; Pred. No. 2.2e+02;
Matches 50; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1209 GCGGCGAGCTGGAGCTGGAGAAGCAGCTGGAGAAGCAGCGGAGCTGGAGCGGCGAGGA 1268
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 81 GCGGCGCTGGAGCTGGAGGAGGAGCGGAGATCACCGGCTGCTGGAGGAGGAGGAC 22
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1269 GGAGGAGGAGGAGGAGAT 1289
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 TGCTGAATCGGAGCTGGAGCT 1
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 9
US-08-556-978B-24
; Sequence 24, Application US/08556978B
; Patent No. 6268169
; GENERAL INFORMATION:
; APPLICANT: FAHNESTOCK, STEPHEN F.
; TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556,978B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,600
; FILING DATE: JUNE 15, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9389-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-556-978B-24

Query Match 0.8%; Score 30.4; DB 3; Length 93;
Best Local Similarity 71.4%; Pred. No. 4.1e+02;
Matches 40; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1708 GCCTTGGAAAGCAAGAGGAGTGGCCGCGGAGCAGCTCCGGGAGGAGCTGGGACGAGGT 1763
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16 GCCTTGGTGTCTAGGCTGCTGGCGGCGGAGCCGCTGGCGGAGCTGGTGGTGGTGGT 71
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-08-556-978B-25/c
; Sequence 25, Application US/08556978B
; Patent No. 6268169
; GENERAL INFORMATION:
; APPLICANT: FAHNESTOCK, STEPHEN F.
; TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
```

```
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556,978B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,600
; FILING DATE: JUNE 15, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9389-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-556-978B-25

Query Match          0.8%; Score 30.4; DB 3; Length 93;
Best Local Similarity 71.4%; Pred. No. 4.1e+02;
Matches 40; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1708 GCCTTGAAGCAAGAGCTGGCGCCGCGACAGCTCGGGAGCAGCTGGACGAGGT 1763
Db 75 GCCTTGGTGTCTCAGGGTGTCTGGCGCGACGCCGCTGCGGCGAGCTGGTGTGCTGT 20

RESULT 11
US-08-674-747-1/c
; Sequence 1, Application US/08068747
; Patent No. 5695933
; GENERAL INFORMATION:
; APPLICANT: Schalling, Martin
; APPLICANT: Hudson, Thomas J.
; APPLICANT: Housman, David E.
; TITLE OF INVENTION: Direct Determination of Expanded
; TITLE OF INVENTION: Nucleotide Repeats in the Human Genome
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,747
; FILING DATE: 28-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-6141
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic"
; US-08-068-747-1

Query Match          0.8%; Score 30.2; DB 1; Length 51;
Best Local Similarity 74.5%; Pred. No. 3.2e+02;
Matches 38; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1213 CAGCTGGAGCTGGAGNAGCAGCTGGAGAGCAGCGGAGCTGGAGCGGCGAG 1263
Db 51 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1

RESULT 12
US-09-513-999C-14585
; Sequence 14585, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14585
; LENGTH: 86
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 40
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 57
; OTHER INFORMATION: r=a or g
; US-09-513-999C-14585

Query Match          0.8%; Score 30; DB 4; Length 86;
Best Local Similarity 59.3%; Pred. No. 5e+02;
Matches 48; Conservative 1; Mismatches 32; Indels 0; Gaps 0;

QY 1207 AAGCGGCGAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAGCGGCGGA 1266
Db 4 AAGCGGCGTGGCGGCTGGAGCAGGAGCGGCGGAGCGGCGGAGCGGCGGCGGGA 63

QY 1267 GAGCAGGAGGAGGAGGAG 1287
Db 64 GCGATGCTGAAGATGGCGGCG 84

RESULT 13
US-07-814-220-26
; Sequence 26, Application US/07814220
; Patent No. 5925540
; GENERAL INFORMATION:
; APPLICANT: Caceci, Thomas
; APPLICANT: Toth, Thomas E.
; APPLICANT: Szumanski, Maria B.W.
; TITLE OF INVENTION: SYNTHETIC ANTIFREEZE PEPTIDE AND
; TITLE OF INVENTION: SYNTHETIC GENE CODING FOR ITS PRODUCTION
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WHITHAM, CURTIS & WHITHAM
; STREET: Reston Int'l. Center, 11800 Sunrise Valley Dr.,
; STREET: Suite 900
```

```
; CITY: Reston
; STATE: VA
; COUNTRY: USA
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/814,220
; FILING DATE: 23-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/588,437
; FILING DATE: 25-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitham, Michael E.
; REGISTRATION NUMBER: 32,635
; REFERENCE/DOCKET NUMBER: CIT.016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-391-2510
; TELEFAX: 703-391-9035
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
; US-07-814-220-26

Query Match 0.8%; Score 29.4; DB 2; Length 57;
Best Local Similarity 70.9%; Pred. No. 5.6e+02;
Matches 39; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1210 CGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAGCGGCGCAGC 1264
Db 1 CCGCAGTTGCTGCCGAGCAGCTGCTGCGGTAGCAGCGGTAGCTGCCGAGCGCGC 55

RESULT 14
US-07-812-421-26
; Sequence 26, Application US/07812421
; Patent No. 5932697
; GENERAL INFORMATION:
; APPLICANT: Caceci, Thomas
; APPLICANT: Toth, Thomas E.
; APPLICANT: Szumanski, Maria B.W.
; TITLE OF INVENTION: SYNTHETIC ANTIFREEZE PEPTIDE AND
; TITLE OF INVENTION: SYNTHETIC GENE CODING FOR ITS PRODUCTION
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WHITHAM, CURTIS & WHITHAM
; STREET: Reston Intl. Center, 11800 Sunrise Valley Dr.,
; STREET: Suite 900
; CITY: Reston
; STATE: VA
; COUNTRY: USA
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/812,421
; FILING DATE: 23-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/588,437
; FILING DATE: 25-SEP-1990
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitham, Michael E.
; REGISTRATION NUMBER: 32,635
; REFERENCE/DOCKET NUMBER: CIT.016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-391-2510
; TELEFAX: 703-391-9035
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
; US-07-812-421-26

Query Match 0.8%; Score 29.4; DB 2; Length 57;
Best Local Similarity 70.9%; Pred. No. 5.6e+02;
Matches 39; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1210 CGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAGCGGCGCAGC 1264
Db 1 CCGCAGTTGCTGCCGAGCAGCTGCTGCGGTAGCAGCGGTAGCTGCCGAGCGCGC 55

RESULT 15
US-07-814-220-25/c
; Sequence 25, Application US/07814220
; Patent No. 5925540
; GENERAL INFORMATION:
; APPLICANT: Caceci, Thomas
; APPLICANT: Toth, Thomas E.
; APPLICANT: Szumanski, Maria B.W.
; TITLE OF INVENTION: SYNTHETIC ANTIFREEZE PEPTIDE AND
; TITLE OF INVENTION: SYNTHETIC GENE CODING FOR ITS PRODUCTION
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WHITHAM, CURTIS & WHITHAM
; STREET: Reston Intl. Center, 11800 Sunrise Valley Dr.,
; STREET: Suite 900
; CITY: Reston
; STATE: VA
; COUNTRY: USA
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/814,220
; FILING DATE: 23-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/588,437
; FILING DATE: 25-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitham, Michael E.
; REGISTRATION NUMBER: 32,635
; REFERENCE/DOCKET NUMBER: CIT.016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-391-2510
; TELEFAX: 703-391-9035
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 62 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
; US-07-814-220-25
```

	Query Match	0.8%;	Score 29.4;	DB 2;	Length 62;
	Best Local Similarity	70.3%;	Pred. No. 5.9e+02;		
	Matches	39;	Conservative	0;	Mismatches 16; Indels 0; Gaps 0;
QY	1210	CGCGCAGCTGAGCTGGAGAGCAGCTGGAGAACGCGGAGCTGGAGCGGCAGC	1264		
Db	62	CCGCGATTGTCGCCGAGCAGCTGCTCGGTAGCAGCGGTAGCTGCCGACGCCGC	8		

Search completed: April 29, 2005, 02:23:45
Job time : 544.672 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 07:33:31 ; Search time 1939.12 Seconds
(without alignments)
11433.739 Million cell updates/sec

Title: US-09-674-237B-2
Perfect score: 3642
Sequence: 1 agggctcagttccacacc.....tggaccacagccagcaatga 3642

Scoring table: IDENTITY NUC
'Gapop 10.0 , Gapext 1.0

Searched: 5642217 seqs, 3043843248 residues

Total number of hits satisfying chosen parameters: 5511450

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48.8	1.3	60	10	US-09-908-975-10766
2	37.6	1.0	69	10	US-09-854-867-463
3	37.6	1.0	69	19	US-10-786-970A-463
4	36.6	1.0	87	18	US-10-333-894A-6
5	34.6	1.0	63	18	US-10-407-818-8
6	33.4	0.9	93	9	US-09-864-761-24680
7	32.4	0.9	87	19	US-10-759-731A-156
8	32.2	0.9	69	10	US-09-854-867-462
9	32.2	0.9	69	19	US-10-786-970A-462
10	31.2	0.9	77	18	US-10-758-307-66
11	31.2	0.9	77	19	US-10-852-797-339
1	48.8	1.3	60	10	US-09-908-975-10766
2	37.6	1.0	69	10	US-09-854-867-463
3	37.6	1.0	69	19	US-10-786-970A-463
4	36.6	1.0	87	18	US-10-333-894A-6
5	34.6	1.0	63	18	US-10-407-818-8
6	33.4	0.9	93	9	US-09-864-761-24680
7	32.4	0.9	87	19	US-10-759-731A-156
8	32.2	0.9	69	10	US-09-854-867-462
9	32.2	0.9	69	19	US-10-786-970A-462
10	31.2	0.9	77	18	US-10-758-307-66
11	31.2	0.9	77	19	US-10-852-797-339

c	12	31.2	0.9	86	14	US-10-085-906-174	Sequence 174, App
c	13	30.8	0.8	99	16	US-10-029-386-17340	Sequence 17340, A
c	14	30.4	0.8	87	19	US-10-759-731A-155	Sequence 155, App
c	15	30.4	0.8	96	16	US-10-061-201-11	Sequence 11, Appl
c	16	29.4	0.8	91	16	US-10-029-386-22222	Sequence 22222, A
c	17	29.4	0.8	81	10	US-09-373-658-52	Sequence 52, Appl
c	18	29.4	0.8	81	11	US-09-989-687-52	Sequence 52, Appl
c	19	29.4	0.8	86	19	US-10-759-731A-158	Sequence 158, App
c	20	29.4	0.8	97	14	US-10-023-066A-57	Sequence 57, Appl
c	21	29.4	0.8	97	18	US-10-804-678-57	Sequence 57, Appl
c	22	29.4	0.8	99	18	US-10-021-323-14338	Sequence 14338, A
c	23	29.2	0.8	87	19	US-10-759-731A-157	Sequence 157, App
c	24	29.2	0.8	99	16	US-10-323-051-21	Sequence 21, Appl
c	25	29	0.8	88	19	US-10-488-936-2	Sequence 2, Appl
c	26	29	0.8	96	18	US-10-021-323-14266	Sequence 14266, A
c	27	28.8	0.8	89	18	US-10-021-323-14335	Sequence 14335, A
c	28	28.8	0.8	89	18	US-10-021-323-15706	Sequence 15706, A
c	29	28.6	0.8	51	18	US-10-865-478-506	Sequence 506, App
c	30	28.2	0.8	58	16	US-10-275-071-15	Sequence 15, Appl
c	31	28.2	0.8	84	14	US-10-023-066A-65	Sequence 65, Appl
c	32	28.2	0.8	84	18	US-10-804-678-65	Sequence 65, Appl
c	33	28	0.8	68	14	US-10-096-986-58	Sequence 58, Appl
c	34	28	0.8	75	18	US-10-645-471A-31	Sequence 31, Appl
c	35	28	0.8	79	17	US-10-023-066A-66	Sequence 2343, Ap
c	36	28	0.8	84	14	US-10-023-066A-66	Sequence 66, Appl
c	37	28	0.8	84	18	US-10-804-678-66	Sequence 66, Appl
c	38	28	0.8	99	17	US-10-072-809A-11	Sequence 11, Appl
c	39	27.8	0.8	88	18	US-10-021-323-808	Sequence 808, App
c	40	27.8	0.8	93	10	US-09-738-937-14	Sequence 14, Appl
c	41	27.8	0.8	96	17	US-10-353-678-66	Sequence 66, Appl
c	42	27.6	0.8	97	9	US-09-864-761-22933	Sequence 22933, A
c	43	27.6	0.8	78	9	US-09-922-261-229	Sequence 229, App
c	44	27.6	0.8	90	17	US-10-296-734-1351	Sequence 1351, Ap
c	45	27.4	0.8	81	18	US-10-021-323-14096	Sequence 14096, A

ALIGNMENTS

RESULT 1

US-09-908-975-10766
; Sequence 10766, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10766
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-10766

Query Match	1.3%	Score 48.8	DB 10	Length 60
Best Local Similarity	88.3%	Pred. No. 0.0022		
Matches 53	Conservative	0	Mismatches 7	Indels 0
Gaps	0			
2231	GAGCGCTGTACCCCTTTGAATCCAGAGTCAGATGATCATCCAGCCAGGAGATA	22990		
1	GGGCACTGTACCCCTTTGAATCCAGAGTCAGATGATCATCCAGCCAGGAGACA	60		

RESULT 2
US-09-854-867-463
; Sequence 463, Application US/09854867
; Publication No. US20030224356A1
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL H
; APPLICANT: ROGAN, PETER K
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/854,867
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 613
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 463
; LENGTH: 69
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: repetitive sequence found in
; OTHER INFORMATION: many eutherian genomes. Length of core repeating element is vari
; OTHER INFORMATION: able and is often polymorphic
US-09-854-867-463
Query Match 1.0%; Score 37.6; DB 10; Length 69;
Best Local Similarity 72.1%; Pred. No. 3.9;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1197 GCAGGAGCCCAAGCGGCGAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGCGAGCTGGA 1256
DB 1 GCA 60
QY 1257 GCGGCAGC 1264
DB 61 GCAGCAGC 68

RESULT 3
US-10-786-970A-463
; Sequence 463, Application US/10786970A
; Publication No. US2005006449A1
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/10/786,970A
; CURRENT FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US/09/573,080
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 463
; LENGTH: 69
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: repetitive sequence found in
; OTHER INFORMATION: many eutherian genomes. Length of core repeating element is vari
; OTHER INFORMATION: able and is often polymorphic
US-10-786-970A-463
Query Match 1.0%; Score 37.6; DB 19; Length 69;
Best Local Similarity 72.1%; Pred. No. 3.9;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1197 GCAGGAGCCCAAGCGGCGAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGCGAGCTGGA 1256
DB 1 GCA 60
QY 1257 GCGGCAGC 1264
DB 61 GCAGCAGC 68

Db 61 GCAGCAGC 68

RESULT 4
US-10-333-894A-6
; Sequence 6, Application US/10333894A
; Publication No. US20040259085A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Chawnshang
; APPLICANT: Hsiung, Ann
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR PREDICTING
; TITLE OF INVENTION: PROSTATE CANCER
; FILE REFERENCE: 21108.0001U1
; CURRENT APPLICATION NUMBER: US/10/333,894A
; CURRENT FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/note =
; OTHER INFORMATION: synthetic construct
US-10-333-894A-6
Query Match 1.0%; Score 36.6; DB 18; Length 87;
Best Local Similarity 65.1%; Pred. No. 8.5;
Matches 54; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 1192 CAGGAGCAGGAGCCCAAGCGGCGAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAG 1251
DB 4 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAGCA 63
QY 1252 CTGGAGCGGCGAGCAGCAGGGA 1274
DB 64 CAGCAACAGCAGCAACAGCAGCA 86

RESULT 5
US-10-407-818-8
; Sequence 8, Application US/10407818
; Publication No. US20040198971A1
; GENERAL INFORMATION:
; APPLICANT: RABBANI, ELAZAR
; APPLICANT: STAVRIANOPOULOS, JANNIS G.
; APPLICANT: DONEGAN, JAMES J.
; TITLE OF INVENTION: MULTISIGNAL LABELING REAGENTS, AND PROCESSES AND USES
; FILE REFERENCE: ENZ-65
; CURRENT APPLICATION NUMBER: US/10/407,818
; CURRENT FILING DATE: 2003-04-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 63
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule:
; OTHER INFORMATION: Synthetic oligonucleotide
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: oligonucleotide
; FEATURE:
; OTHER INFORMATION: 3'-amided
US-10-407-818-8
Query Match 1.0%; Score 34.6; DB 18; Length 63;
Best Local Similarity 75.4%; Pred. No. 27;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1208 AGCGGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAGCGGAGC 1264


```

RESULT 7
US-10-759-731A-156/c
; Sequence 156, Application US/10759731A
; Publication No. US20050079574A1
; GENERAL INFORMATION:
; APPLICANT: Bond, Christopher J.
; TITLE OF INVENTION: SYNTHETIC ANTIBODY PHAGE LIBRARIES
; FILE REFERENCE: 11669.136USU1
; CURRENT APPLICATION NUMBER: US/10/759,731A
; CURRENT FILING DATE: 2004-01-16
; PRIOR APPLICATION NUMBER: US 60/441,059
; PRIOR FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US 60/488,610
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/510,314
; PRIOR FILING DATE: 2003-10-08
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 156
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: VLX ala scan
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (26)..(26)
; OTHER INFORMATION: y is c o r t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28)..(28)
; OTHER INFORMATION: s is g o r c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: y is c o r t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: r is a o r g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: m is a o r c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (34)..(35)
; OTHER INFORMATION: s is g o r c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (37)..(38)
; OTHER INFORMATION: s is g o r c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: s is g o r c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (43)..(43)
; OTHER INFORMATION: k is g o r t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (46)..(46)

```


; APPLICANT: Shak, Steven
 ; APPLICANT: Baker, Joffe
 ; APPLICANT: Cronin, Maureen
 ; TITLE OF INVENTION: GENE EXPRESSION MARKERS FOR BREAST
 ; FILE REFERENCE: 39740/0008 US
 ; CURRENT APPLICATION NUMBER: US/10/758,307
 ; PRIOR FILING DATE: 2004-01-14
 ; PRIOR APPLICATION NUMBER: US 60/440,861
 ; PRIOR FILING DATE: 2003-01-15
 ; NUMBER OF SEQ ID NOS: 440
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 66
 ; LENGTH: 77
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Amplicon
 US-10-758-307-66

Query Match 0.9%; Score 31.2; DB 18; Length 77;
 Best Local Similarity 70.0%; Pred. No. 2.8e+02;
 Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
 QY 1186 GAGCGCCAGGAGCAGGAGGCGGCGGCTGAGCTGGAGAGCAGCTGGAGAGCAG 1245
 DB 2 GAGCGCGAGATCAGGAGTACCGCGGCTCATGGACATCAAGTCGCGGCTGGAGCAGGAG 61

RESULT 11
 US-10-852-797-339
 ; Sequence 339, Application US/10852797
 ; Publication No. US20050064455A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Genomic Health, Inc.
 ; APPLICANT: Baker, Joffe
 ; APPLICANT: Miller, Kathy D.
 ; APPLICANT: Shak, Steven
 ; APPLICANT: Sledge, George
 ; APPLICANT: Soule, Sharon
 ; TITLE OF INVENTION: Gene Expression Markers for Predicting
 ; TITLE OF INVENTION: Response to Chemotherapy
 ; FILE REFERENCE: 39740-0010
 ; CURRENT APPLICATION NUMBER: US/10/852,797
 ; CURRENT FILING DATE: 2004-05-24
 ; PRIOR APPLICATION NUMBER: 60/473,970
 ; PRIOR FILING DATE: 2003-05-28
 ; NUMBER OF SEQ ID NOS: 372
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 339
 ; LENGTH: 77
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: amplicon
 US-10-852-797-339

Query Match 0.9%; Score 31.2; DB 19; Length 77;
 Best Local Similarity 70.0%; Pred. No. 2.8e+02;
 Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
 QY 1186 GAGCGCCAGGAGCAGGAGGCGGCGGCTGAGCTGGAGAGCAGCTGGAGAGCAG 1245
 DB 2 GAGCGCGAGATCAGGAGTACCGCGGCTCATGGACATCAAGTCGCGGCTGGAGCAGGAG 61

RESULT 12
 US-10-085-906-174/c
 ; Sequence 174, Application US/10085906
 ; Publication No. US20030054371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ying, Vincent
 ; APPLICANT: Wu, Paul

; APPLICANT: Gray, Gary S.
 ; TITLE OF INVENTION: POLYMORPHIC ELEMENTS IN THE
 ; TITLE OF INVENTION: COSTIMULATORY RECEPTOR LOCUS AND USES THEREOF
 ; FILE REFERENCE: GNN-5343CP2
 ; CURRENT APPLICATION NUMBER: US/10/085,906
 ; CURRENT FILING DATE: 2002-02-27
 ; PRIOR APPLICATION NUMBER: US 60/126,215
 ; PRIOR FILING DATE: 1999-03-25
 ; PRIOR APPLICATION NUMBER: US 09/534,061
 ; PRIOR FILING DATE: 2000-03-24
 ; PRIOR APPLICATION NUMBER: PCT/US00/07938
 ; PRIOR FILING DATE: 2000-03-24
 ; NUMBER OF SEQ ID NOS: 545
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 174
 ; LENGTH: 86
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-085-906-174

Query Match 0.9%; Score 31.2; DB 14; Length 86;
 Best Local Similarity 60.7%; Pred. No. 3e+02;
 Matches 51; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
 QY 1219 GAGCTGGAGAACGAGCTGGAGAGCAGCGGGAGCTGGAGCGGCGGAGGAGG 1278
 DB 85 GAGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 26
 QY 1279 AGGAGAGGAGATCGAGGCGGCGAG 1302
 DB 25 AAGGAGAGAGGAGGAGGAGGAGGAG 2

RESULT 13
 US-10-029-386-17340
 ; Sequence 17340, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
 ; FILE REFERENCE: AEOMICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029,386
 ; CURRENT FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 17340
 ; LENGTH: 99
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL135920.4
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 8.7
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 13
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 8
 ; OTHER INFORMATION: NT HIT: D86240.2, EVALUATE 4.90e+00
 US-10-029-386-17340

Query Match 0.8%; Score 30.8; DB 16; Length 99;
 Best Local Similarity 63.5%; Pred. No. 4.2e+02;
 Matches 47; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
 QY 1169 AGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1228
 DB 11 AGAAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 70
 QY 1229 AGCAGCTGGAGAG 1242
 DB 71 AGGAGAGGAGGAG 84

```
RESULT 14
US-10-759-731A-155/c
; Sequence 155, Application US/10759731A
; Publication No. US20050079574A1
; GENERAL INFORMATION:
; APPLICANT: Bond, Christopher J.
; TITLE OF INVENTION: SYNTHETIC ANTIBODY PHAGE LIBRARIES
; FILE REFERENCE: 11669.136USU1
; CURRENT APPLICATION NUMBER: US/10759,731A
; CURRENT FILING DATE: 2004-01-16
; PRIOR APPLICATION NUMBER: US 60/441,059
; PRIOR FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US 60/489,610
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/510,314
; PRIOR FILING DATE: 2003-10-08
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 155
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: RIG ala scan
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (25)..(26)
; OTHER INFORMATION: s is g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28)..(28)
; OTHER INFORMATION: r is a or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: y is c or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: s is g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (34)..(35)
; OTHER INFORMATION: s is g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (37)..(37)
; OTHER INFORMATION: k is g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: y is c or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (43)..(43)
; OTHER INFORMATION: k is g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44)..(44)
; OTHER INFORMATION: y is c or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (46)..(46)
; OTHER INFORMATION: r is a or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: m is a or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (49)..(49)
; OTHER INFORMATION: s is g or c
```

```
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (50)..(50)
; OTHER INFORMATION: y is c or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (52)..(53)
; OTHER INFORMATION: s is g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55)..(56)
; OTHER INFORMATION: s is g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (59)..(59)
; OTHER INFORMATION: m is a or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (61)..(61)
; OTHER INFORMATION: k is g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (64)..(64)
; OTHER INFORMATION: k is g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65)..(65)
; OTHER INFORMATION: s is g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (68)..(68)
; OTHER INFORMATION: y is c or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (70)..(70)
; OTHER INFORMATION: r is a or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (73)..(73)
; OTHER INFORMATION: k is g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (74)..(74)
; OTHER INFORMATION: s is g or c
US-10-759-731A-155

Query Match      0.8%; Score 30.4; DB 19; Length 87;
Best Local Similarity 39.5%; Pred.No.5.le+02;
Matches 34; Conservative 23; Mismatches 29; Indels 0; Gaps 0;

QY 1196 AGCAGAGGCCAAGCGGCGAGCTGGAGAGCAGCTGGAGAGCAGCTGGAGAGCAGCTGGAGAGCAGCTGGAGCTGG 1255
Db 87 ACCTGACCCACACSMAGYARCCSMGMGMYKCASSASSARSRGKYARWARGCGMASSASCARY 28
QY 1256 AGCGGAGCGGAGGAGGAGGAGGAGG 1281
Db 27 ASSACCAGCACCACCAAGTATAGACGG 2

RESULT 15
US-10-061-201-11
; Sequence 11, Application US/10061201
; Publication No. US20030166229A1
; GENERAL INFORMATION:
; APPLICANT: Shannon, Mark
; TITLE OF INVENTION: HUMAN POSH-LIKE PROTEIN 1
; FILE REFERENCE: PB0178
; CURRENT APPLICATION NUMBER: US/10/061,201
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 28, 2005, 21:36:20 ; Search time 12119 Seconds
(without alignments)
4849.920 Million cell updates/sec

Title: US-09-674-237B-3

Perfect score: 6269

Sequence: 1 MQAFPFGSLDVAITVE.....VGLFPSNYKLTMDPSQ 1213

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 2238514

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09674237/runat.27042005.182121.2876/app.query.fasta_1.1351
-DB=GenEmbl -QFMT=fastap -SUFFIX=szml100.rge -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=100
-USER=US09674237@cgn.1.1.7319 @runat.27042005.182121.2876 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	1.6	60	6	CQ541131 Sequence
2	88	1.4	96	6	AR208350 Sequence
3	88	1.4	96	6	AR208351 Sequence
4	79.5	1.3	98	3	AY545898 Fennerope

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	76.5	1.2	91	8	AF263002	AF263002 Ditylum b
c	75.5	1.2	99	3	AY184035	AY184035 Drosophil
c	75.5	1.2	99	3	AY184060	AY184060 Drosophil
c	75.5	1.2	100	3	AY184027	AY184027 Drosophil
c	74.5	1.2	99	3	AY184026	AY184026 Drosophil
c	74.5	1.2	90	3	AY545903	AY545903 Fennerope
c	74	1.2	93	8	CMY17097	Y17097 Cladobotryu
c	73	1.2	93	8	AY545901	AY545901 Fennerope
c	72	1.1	98	3	AY545901	AY545901 Fennerope
c	71	1.1	78	8	AY293995	AY293995 Arabidops
c	71	1.1	90	8	AY295846	AY295846 Arabidops
c	70.5	1.1	77	8	AY293994	AY293994 Arabidops
c	67.5	1.1	96	6	AX530502	AX530502 Sequence
c	67.5	1.1	96	6	AX530502	AX530502 Sequence
c	66	1.1	79	3	AY184038	AY184038 Drosophil
c	66	1.1	79	3	AY184047	AY184047 Drosophil
c	66	1.1	79	3	AY184053	AY184053 Drosophil
c	66	1.1	84	3	AY184051	AY184051 Drosophil
c	66	1.1	85	3	AY132812	AY132812 Fennerope
c	65.5	1.0	87	6	AX020306	AX020306 Sequence
c	65.5	1.0	87	6	BD037839	BD037839 Sequence
c	65	1.0	83	10	MMVIMV30	X89147 M.musculus
c	64.5	1.0	73	8	AY293993	AY293993 Arabidops
c	64	1.0	82	6	AR421455	AR421455 Sequence
c	64	1.0	82	6	AX982149	AX982149 Sequence
c	64	1.0	82	6	BD117008	BD117008 EST and e
c	63.5	1.0	96	6	AR107476	AR107476 Sequence
c	63	1.0	60	6	CQ542573	CQ542573 Sequence
c	62.5	1.0	86	8	AY295845	AY295845 Arabidops
c	61.5	1.0	60	6	CQ538403	CQ538403 Sequence
c	61.5	1.0	99	6	CQ725585	CQ725585 Sequence
c	61.5	1.0	100	4	AY045524	AY045524 Panthera
c	61	1.0	75	3	AY184049	AY184049 Drosophil
c	61	1.0	79	3	AY184031	AY184031 Drosophil
c	61	1.0	94	11	HUMUT1531B	L30884 Human STS U
c	60	1.0	58	6	AR208349	AR208349 Sequence
c	60	1.0	67	8	AY293992	AY293992 Arabidops
c	60	1.0	100	6	A43830	AY293992 Arabidops
c	60	1.0	100	6	I17407	I17407 Sequence 4
c	59.5	0.9	98	6	BD270475	BD270475 Synthetic
c	59.5	0.9	98	6	AX039308	AX039308 Sequence
c	59.5	0.9	98	6	AX039517	AX039517 Sequence
c	59.5	0.9	100	3	AF411993	AF411993 Formica e

ALIGNMENTS

RESULT 1	CQ541131	Sequence 10766 from Patent WO210449.	60 bp	DNA	linear	PAT 30-JAN-2004
CQ541131	LOCUS	CQ541131				
	DEFINITION	CQ541131				
	ACCESSION	CQ541131				
	VERSION	CQ541131.1	GI:41507395			
	KEYWORDS					
	SOURCE	Homo sapiens (human)				
	ORGANISM	Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.				
	REFERENCE	1				
	AUTHORS	Shoshan,A., Wasserman,A., Mintz,E., Mintz,L. and Faigler,S.				
	TITLE	Oligonucleotide library for detecting rna transcripts and splice variants that populate a transcriptome				
	JOURNAL	Patent: WO 0210449-A 10766 07-FEB-2002;				
		CompuGen Inc. (US)				
	FEATURES	Location/Qualifiers				
	source	1..60				
		/organism="Homo sapiens"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:9606"				

ORIGIN

Alignment Scores:	3.49e+03	Length:	60
Pred. No.:	102.00	Matches:	19
Score:	100.00%	Conservative:	0

```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.63% Indels: 0
DB: 6 Gaps: 0

US-09-674-237B-3 (1-1213) x CQ541131 (1-60)

QY 745 AlaLeuTyrProPheGluSerArgSerHisAspGluLeuThrIleGlnProGlyAsp 763
DB 3 GCACTGTACCCCTTGATTCAGAGCCATGATGAATCACTATCCAGCCAGAGAC 59

RESULT 2
LOCUS AR208350/c
DEFINITION Sequence 6 from patent US 6383747.
ACCESSION AR208350
VERSION AR208350.1 GI:21509481
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 96)
AUTHORS Dawkins,R.Letts, and Abraham,L.Joseph.
TITLE Method for determining ancestral haplotypes using haplospesific
geometric elements within the major histocompatibility complex
multigene cluster
JOURNAL Patent: US 6383747-A 6 07-MAY-2002;
FEATURES Location/Qualifiers
source 1..96
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.94e+04 Length: 96
Score: 88.00 Matches: 18
Percent Similarity: 76.67% Conservatave: 5
Best Local Similarity: 60.00% Mismatches: 7
Query Match: 1.40% Indels: 0
Gaps: 6

US-09-674-237B-3 (1-1213) x AR208350 (1-96)

QY 413 GlnLysGlnArgGluLeuArgGlnArgGluGluGluArgLysGluLeuGluArg 432
DB 92 GAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 33

RESULT 3
LOCUS AR208351/c
DEFINITION Sequence 7 from patent US 6383747.
ACCESSION AR208351
VERSION AR208351.1 GI:21509483
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 96)
AUTHORS Dawkins,R.Letts, and Abraham,L.Joseph.
TITLE Method for determining ancestral haplotypes using haplospesific
geometric elements within the major histocompatibility complex
multigene cluster
JOURNAL Patent: US 6383747-A 7 07-MAY-2002;
FEATURES Location/Qualifiers
source 1..96
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.94e+04 Length: 96
Score: 88.00 Matches: 18
Percent Similarity: 76.67% Conservatave: 5
Best Local Similarity: 60.00% Mismatches: 7
Query Match: 1.40% Indels: 0
Gaps: 6

US-09-674-237B-3 (1-1213) x AR208350 (1-96)

QY 413 GlnLysGlnArgGluLeuArgGlnArgGluGluGluArgLysGluLeuGluArg 432
DB 92 GAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 33

QY 433 ArgGluAlaAlaLysArgGluLeuGluArg 442
DB 32 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3

RESULT 4
LOCUS AY545898/c
DEFINITION Fenneropenaeus chinensis clone TE1-122696 microsatellite sequence.
ACCESSION AY545898
VERSION AY545898.1 GI:44894761
KEYWORDS
SOURCE Fenneropenaeus chinensis
ORGANISM Fenneropenaeus chinensis
Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
Penaeidae; Fenneropenaeus.
REFERENCE 1 (bases 1 to 98)
AUTHORS Kong,J., Meng,X., Liu,P. and Gao,H.
TITLE Selection of polymorphic microsatellites from random shearing DNA
clones in Fenneropenaeus chinensis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 98)
AUTHORS Kong,J., Meng,X., Liu,P. and Gao,H.
TITLE Direct Submission
JOURNAL Submitted (07-FEB-2004) Genoplasm Resources and Genetic Breeding
Laboratory, Yellow Sea Fisheries Research Institute, 106 Nanjing
Road, Qingdao, Shandong 266071, P. R. China

FEATURES
source 1..98
/organism="Fenneropenaeus chinensis"
/mol_type="genomic DNA"
/db_xref="taxon:139456"
/clone="TE1-122696"
repeat_region 1..98
/note="microsatellite"
/rpt_type=tandem

ORIGIN
Alignment Scores:
Pred. No.: 4.11e+04 Length: 98
Score: 79.50 Matches: 17
Percent Similarity: 64.86% Conservatave: 7
Best Local Similarity: 45.95% Mismatches: 6
Query Match: 1.27% Indels: 7
Gaps: 3

US-09-674-237B-3 (1-1213) x AY545898 (1-98)

QY 393 LysGluArgGluArgGlnGlnGlnGluAlaLysArgGlnLeuGluLeuGluLysGlnLeu 412
DB 90 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 52

QY 413 GlnLysGlnArgGluLeuGluArgGlnArgGluGluGluArgLysGlu 429
DB 51 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1

RESULT 5
LOCUS AF263002
DEFINITION Ditylum brightwellii clone Dbr2 microsatellite sequence.
```


ORIGIN

FEATURES
 source
 Location/Qualifiers
 1..93
 /organism="Hypomyces odoratus"
 /mol_type="genomic DNA"
 /isolate="Z15001"
 /specific_host="Agaricus bisporus"
 /db_xref="taxon:76883"
 /tissue_type="mycelium"
 1..93
 /note="microsatellite"

satellite

ORIGIN

Alignment Scores:
 Pred. No.: 6.78e+04 Length: 93
 Score: 73.00 Matches: 13
 Percent Similarity: 76.92% Conservatively: 7
 Best Local Similarity: 50.00% Mismatches: 6
 Query Match: 1.16% Indels: 0
 DB: 8 Gaps: 0

US-09-674-237B-3 (1-1213) x CMY17097 (1-93)

Qy 404 ArgGlnLeuGluLeuGluGlnLeuGluGlnArgGluLeuGluArgGlnArgGlu 423

Db 9 CGCATATGGAG 68

Qy 424 GluGluArgArgLysGlu 429

Db 69 AGAGAGAGAGAGAGAA 86

RESULT 12
 AY545901/c
 LOCUS
 Penneropenaeus chinensis clone TE4-32234 microsatellite sequence.
 ACCESSION
 AY545901
 VERSION
 AY545901.1 GI:44894764
 KEYWORDS
 SOURCE
 ORGANISM
 Penneropenaeus chinensis
 Penneropenaeus chinensis
 Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
 Penaeidae; Penneropenaeus.
 1 (bases 1 to 98)
 Kong, J., Meng, X., Liu, P. and Gao, H.
 Selection of polymorphic microsatellites from random shearing DNA
 clones in Penneropenaeus chinensis
 Unpublished
 2 (bases 1 to 98)
 Kong, J., Meng, X., Liu, P. and Gao, H.
 Direct Submission
 Submitted (07-FEB-2004) Genplasm Resources and Genetic Breeding
 Laboratory, Yellow Sea Fisheries Research Institute, 106 Nanjing
 Road, Qingdao, Shandong 266071, P. R. China
 Location/Qualifiers
 1..98
 /organism="Penneropenaeus chinensis"
 /mol_type="genomic DNA"
 /db_xref="taxon:139456"
 /clone="TE4-32234"
 1..98
 /note="microsatellite"
 /rpt_type=tandem

repeat_region

ORIGIN

Alignment Scores:
 Pred. No.: 7.82e+04 Length: 98
 Score: 72.00 Matches: 13
 Percent Similarity: 75.00% Conservatively: 8
 Best Local Similarity: 46.43% Mismatches: 7
 Query Match: 1.15% Indels: 0
 DB: 3 Gaps: 0

US-09-674-237B-3 (1-1213) x AY545901 (1-98)

```

REFERENCE 1 (bases 1 to 90)
AUTHORS Symonds,V.V. and Lloyd,A.M.
TITLE An Analysis of Microsatellite Loci in Arabidopsis thaliana.
JOURNAL Mutational dynamics and application
GENETICS 165 (3), 1475-1488 (2003)
PUBMED 14668396
REFERENCE 2 (bases 1 to 90)
AUTHORS Symonds,V.V. and Lloyd,A.M.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-2003) MCDB, University of Texas-Austin, 2500
Speedway, Austin, TX 78712, USA
FEATURES
source
1..90
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/ecotype="Pa-3"
repeat_region 1..90
/note="microsatellite ngal72"
/rpt_type=tandem
ORIGIN
Alignment Scores:
Pred. No.:      8.1e+04      Length:      90
Score:          70.50       Matches:    17
Percent Similarity: 53.8%     Conservative: 4
Best Local Similarity: 43.5%   Mismatches: 3
Query Match:    1.12%        Indels:     15
DB:             8           Gaps:       1
US-09-674-237B-3 (1-1213) x AY295846 (1-90)
Qy 388 AlaGluGlnGluArgLysGluArgGluGlnGlnGlnGlnGluAlaLysArgGlnLeuGlu 407
Db 18 GCAGAGAGACCGAAGAAGAGAGAGAGA----- 47
Qy 408 LeuGluLysGlnLeuGlnLysGlnArgGlnLeuGluArgGlnGlnGluGlnArg 426
Db 48 -----GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAACGA 89
RESULT 15
LOCUS AY293994
DEFINITION Arabidopsis thaliana Ob-3 microsatellite nga8 sequence.
ACCESSION AY293994
VERSION AY293994.1 GI:32478850
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
REFERENCE 1 (bases 1 to 77)
AUTHORS Symonds,V.V. and Lloyd,A.M.
TITLE An Analysis of Microsatellite Loci in Arabidopsis thaliana.
JOURNAL Mutational dynamics and application
PUBMED 14668396
REFERENCE 2 (bases 1 to 77)
AUTHORS Symonds,V.V. and Lloyd,A.M.
TITLE Direct Submission
JOURNAL Submitted (08-MAY-2003) MCDB, University of Texas-Austin, 2500
Speedway, Austin, TX 78712, USA
FEATURES
source
1..77
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/ecotype="Ob-3"
repeat_region 1..77
/note="microsatellite nga8"
/rpt_type=tandem
ORIGIN

```

Search completed: April 29, 2005, 10:38:23
Job time : 12126 secs

THIS PAGE BLANK (USPTO)

PT gene of different organism immobilized on support.
XX Disclosure; Page 16; 56pp; Japanese.
XX The invention relates to a novel oligonucleotide array comprising a
CC number of oligonucleotides derived from an orthologue gene of a different
CC organism or species immobilised on a support body. The oligonucleotide
CC array has two sides comprising a gene derived from two different
CC organisms, in which one side comprises a human gene. In the
CC oligonucleotide array a base sequence differs in a different position on
CC the support body. Each of the oligonucleotides is a partial sequence of
CC the orthologue gene of the organism A and organism B. The sequence match
CC degree of the organism species A and the organism species B is less than
CC 70%. The sequence homology of the genes other than the orthologue gene of
CC the organism species A and the organism species B is the level-of-
CC statistical-significance value, calculated using the homology searching
CC algorithm: basic local alignment search tool (BLAST) and the value is 0.1
CC or more. The found value or the calculated value of the melting-
CC temperature of the variant genes other than the orthologue gene of the
CC organism species A and the organism species B is 20 degrees C or more.
CC The oligonucleotide array is useful for measuring expression distribution
CC of the orthologue gene in different organisms. The method is useful for
CC comparing expression change of the orthologue gene with respect to
CC medical-agent administration and for comparing change of the function of
CC the orthologue gene in different organisms. The oligonucleotide array has
CC the ability to perform a measurement of gene-expression distribution of
CC two or more types of organism simultaneously. The oligonucleotide array
CC improves reliability of measurement. This polynucleotide sequence
CC represents a probe of a rat gene for comparison against a human gene used
CC in the oligonucleotide array of the invention.
XX
SQ Sequence 80 BP; 19 A; 15 C; 25 G; 21 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.84 Length: 80
Score: 155.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.47% Indels: 0
DB: 12 Gaps: 0

US-09-674-237B-3 (1-1213) x ADP49483 (1-80)

QY 940 AspMetTrrpPheGlyGluValGlnGlyGlnLysGlyTrrpPheProLysSerTyrVal 959
DB 2 GACATGGTGGTGGTGGAGAGTTCAGGTCAGAGGGTGGTCCCAAGTCTTACGTG 61

QY 960 LysLeuIleSerGlyPro 965
DB 62 AAATCATTTTCAGGGCCC 79

RESULT 2
ID ABN38018 standard; DNA; 60 BP.
XX ABN38018;
XX
XX 15-JUL-2002 (first entry)
XX Human spliced transcript detection oligonucleotide SEQ ID NO:10766.
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Homo sapiens.
XX
XX WO200210449-A2.
XX
XX 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-IB001903.
XX
XX 28-JUL-2000; 2000US-0221607P.
XX
XX

PR 02-MAY-2001; 2001US-0287724P.
XX (COMP-) COMPUGEN INC.
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of a
PT genome, useful for detecting tissue-, pathology-, and developmental-
PT specific genes.
XX
XX Example 1; SEQ ID NO 10766; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
CC)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises several
CC oligonucleotides, each capable of hybridising selectively to a set of
CC messenger RNAs transcribed from a given transcription unit of the genome,
CC which encodes one or more messenger RNA splice variants. The
CC oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a particular
CC biological or pathological state, and so allowing the detection of tissue
CC - and pathology-specific genes such as those genes only expressed in
CC specific tissue under a specific pathological condition; to detect
CC developmental specific genes; and to detect RNA transcripts and splice
CC variants of a transcriptome of a patient suffering from a particular
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC rats, humans and mice, which are used in the exemplification of the
CC present invention. N.B. The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 60 BP; 19 A; 17 C; 13 G; 11 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 644 Length: 60
Score: 102.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.63% Indels: 0
DB: 6 Gaps: 0

US-09-674-237B-3 (1-1213) x ABN38018 (1-60)

QY 745 AlaLeuTyrProPheGluSerArgSerHisAspGluIleThrIleGlnProGlyAsp 763
DB 3 GCACTGTACCCCTTTCATCCAGAGCCATGATGAATCACTATCCAGCCAGGAGAC 59

RESULT 3
ID ACN59557/c
XX ACN59557 standard; cDNA; 99 BP.
XX ACN59557;
XX
XX 02-DEC-2004 (first entry)
XX
XX Cotton gynoeicum tissue EST Clone ID: LIB3829-032-Q6-N6-F4, SEQ:14338.
XX
XX Cotton; plant; EST; expressed sequence tag; transgenic plant; gynoeicum;
KW variety Nu cotton33B; library LIB3829; molecular tag; molecular marker;
KW genetic mapping; molecular mapping; seed germination; plant growth;
KW plant quality; plant yield; plant breeding; tissue printing; ss.
XX
XX Gossypium hirsutum.
XX
XX US2004123340-A1.
XX
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 28, 2005, 20:56:09 ; Search time 1397 Seconds

(without alignments)
5140.047 Million cell updates/sec

Title: US-09-674-237B-3

Perfect score: 6269

Sequence: 1 MAQFPFPFGSLDVAITVE.....VGLFPSNYVKLTMDPSQQ 1213

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 4530610

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh
-Q=/cpn2_1/USPTO.spool/US09674237/runat_27042005.182121.2866/app.query.fasta.1.1351
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFFIX=szml100.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=100
-USER=US09674237@CGN.1.1.717@runat_27042005.182121.2866 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	155	2.5	80	ADP49483	Adp49483 Oligonucleotide array related rat oligonucleotide probe No 99.
2	102	1.6	60	ABN38018	Abn38018 Human spl
3	89	1.4	99	ACN59557	Acn59557 Cotton gy
4	84	1.3	96	ACN59485	Acn59485 Cotton gy
5	79	1.3	88	ACN46027	Acn46027 Cotton pr

C	6	79	1.3	89	13	ACN60925	Acn60925 Cotton gy
C	7	79	1.3	89	13	ACN59554	Acn59554 Cotton gy
C	8	77	1.2	88	8	ACC69626	Acc69626 Peptide n
C	9	73	1.2	81	13	ACN59315	Acn59315 Cotton gy
C	10	73	1.2	98	4	AAI84741	AAI84741 Human pol
C	11	70	1.1	73	13	ACN59545	Acn59545 Cotton gy
C	12	70	1.1	76	13	ACN61067	Acn61067 Cotton gy
C	13	67.5	1.1	96	6	ABV89298	Abv89298 Human POS
C	14	66	1.1	98	12	ADJ36166	Adj36166 Self-coal
C	15	65.5	1.0	87	3	AAC14094	Aac14094 Human sec
C	16	65	1.0	69	7	ADS31416	Ads31416 Human gen
C	17	64.5	1.0	95	2	AAT23132	Aat23132 Human gen
C	18	64	1.0	73	12	ADH70530	Adh70530 Human Vbe
C	19	64	1.0	96	4	AAH84246	Aah84246 Human cel
C	20	63.5	1.0	96	2	AAT49258	Aat49258 HVR1 regi
C	21	63	1.0	60	6	ABN39460	Abn39460 Human spl
C	22	62	1.0	66	4	AAI84742	AAI84742 Human pol
C	23	61.5	1.0	60	6	ABN35290	Abn35290 Human spl
C	24	60.5	1.0	73	4	AAI84743	AAI84743 Human pol
C	25	60	1.0	92	12	ACH89288	Ach89288 Human gen
C	26	60	1.0	100	2	AAQ62560	Aaq62560 Candida a
C	27	59.5	0.9	98	4	AAC65404	Aac65404 Oligonucle
C	28	59.5	0.9	100	4	AAC65405	Aac65405 Oligonucle
C	29	59	0.9	77	3	AAF18418	Aaf18418 Lung canc
C	30	58	0.9	58	2	AAQ33612	Aaq33612 Microsate
C	31	58	0.9	59	13	ACN46043	Acn46043 Cotton pr
C	32	58	0.9	60	13	ACN59546	Acn59546 Cotton gy
C	33	58	0.9	67	4	AAI84826	AAI84826 Human pol
C	34	58	0.9	74	12	ADL60206	Adl60206 Rat cdna
C	35	58	0.9	78	12	ADJ36167	Adj36167 Self-coal
C	36	58	0.9	90	12	ACH88910	Ach88910 Human gen
C	37	58	0.9	92	4	AAC65406	Aac65406 Oligonucle
C	38	58	0.9	94	4	AAC65407	Aac65407 Oligonucle
C	39	58	0.9	96	2	AAQ33526	Aaq33526 Sequence
C	40	58	0.9	96	3	AAZ57144	Aaz57144 Exemplary
C	41	58	0.9	98	12	ACH89206	Ach89206 Human gen
C	42	58	0.9	100	2	AAQ62559	Aaq62559 Candida a
C	43	57.5	0.9	96	3	AAZ57135	Aaz57135 Quadruple
C	44	57	0.9	65	6	ABN57047	Abn57047 Mouse spl
C	45	57	0.9	65	6	ABN54472	Abn54472 Mouse spl

ALIGNMENTS

RESULT 1
ADP49483
ID ADP49483 standard; DNA; 80 BP.
XX
AC ADP49483;
DT 12-AUG-2004 (first entry)
XX
XX Oligonucleotide array related rat oligonucleotide probe No 99.
XX oligonucleotide array; orthologue; homology; expression distribution;
KW change; gene-expression; rat; probe; ss.
XX Rattus norvegicus.
XX
XX JP2004016070-A.
XX
XX 22-JAN-2004.
XX
XX 14-JUN-2002; 2002JP-00174208.
XX
XX 14-JUN-2002; 2002JP-00174208.
XX
XX (HITA) HITACHI LTD.
PA
XX
XX WPI; 2004-113862/12.
XX
XX Oligonucleotide array, useful for measuring ortholog gene-expression
PT distribution, comprising number of oligonucleotides derived from ortholog

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 28, 2005, 22:12:55 ; Search time 8531 Seconds
(without alignments)
5412.256 Million cell updates/sec

Title: US-09-674-237B-3
Perfect score: 6269
Sequence: 1 MQAFPFGSLDVAITVE.....VGLFPSNYKLTMDMPDPSQQ 1213

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 675282

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cpn2.1/USPTO.spool/US09674237/runat.27042005.182122.2888/app.query.fasta.1.1351
-DB=EST -QPM=fastap -SUFFIX=szml100.rst -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=b1cs -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=100
-USER=US09674237 @CGN 1.1 5042 @runat.27042005.182122.2888 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DBLEXT=7

Database: EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_hic.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_gss1.*
- 9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	168	2.7	94	8	CC199586 XE464 Bay
2	122	1.9	79	1	AA589590 v149C09.s
3	95.5	1.5	99	2	AV963382 AV963382
4	89	1.4	99	7	CK891128 SGP163390
5	87.5	1.4	98	6	CB262655 53-E8866-
6	87	1.4	96	1	AL821665 AL821665
7	87	1.4	97	1	AU250968 AU250968
8	87	1.4	97	2	AW691901 NF050E01S
9	87	1.4	97	2	AW693216 NF061F12S

C	10	87	1.4	97	6	CB256469	77-E01084
C	11	87	1.4	97	7	CR580027	CR580027
	12	87	1.4	98	7	CK878555	SGP139924
	13	86	1.4	90	7	CO909028	BJ02039B0
	14	86	1.4	100	7	CV295974	EST884351
	15	85	1.4	90	7	CO912202	BJ03029F0
	16	84	1.3	97	1	AU246145	AU246145
	17	84	1.3	98	6	CB886842	Lice_O1_N
	18	83.5	1.3	100	7	CK459729	929944 MA
	19	83	1.3	102	7	CV296344	EST884721
C	20	83	1.3	96	6	BE662207	bs31G06.Y
	21	83	1.3	98	6	CB224209	1JEJ32F6
	22	83	1.3	100	7	CK999753	io89f03.b
	23	83	1.3	100	9	CR135046	Reverse s
	24	82	1.3	93	5	BX255382	BX255382
	25	82	1.3	93	6	CB922319	VVD082C06
	26	82	1.3	94	2	BF645175	NF033F01E
	27	82	1.3	94	2	BE402129	CSB004G03
	28	82	1.3	94	5	BQ607527	BRY_3421
C	29	82	1.3	95	6	CB257825	45-E01084
C	30	82	1.3	95	7	CR438601	CR438601
	31	82	1.3	97	1	AI947161	bs35G07.Y
	32	82	1.3	98	6	CD487616	Gm_ckt295
	33	81.5	1.3	96	2	BE976171	bs47G02.Y
	34	81.5	1.3	98	7	CO912138	BJ03028H0
	35	81.5	1.3	100	6	C21938	C21938 Miya
	36	81	1.3	95	7	CV296657	EST885034
	37	80	1.3	86	7	CV297044	EST885421
	38	80	1.3	90	7	CO911413	BJ03018H0
	39	80	1.3	91	2	AW693933	NF070G02S
	40	80	1.3	91	6	CB005813	VVC026C03
C	41	80	1.3	91	7	CR432141	CR432141
	42	80	1.3	92	2	BE317674	NF053E02L
C	43	80	1.3	92	6	CB918229	VVD030C08
	44	80	1.3	92	7	CO907376	BJ02017A0
	45	80	1.3	92	7	CO907631	BJ02019H0

ALIGNMENTS

RESULT 1
CC199586

LOCUS
DEFINITION

CC199586 XE464 BayGenomics Gene Trap Library pGTL1xf Mus musculus cDNA, mRNA

sequence.

CC199586 CC199586.1 GI:30479626

KEYWORDS

SOURCE GSS.

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 94)

BayGenomics.

http://baygenomics.ucsf.edu/

Unpublished (2001)

CONTACT: BayGenomics

Bay Area Functional Genomics Consortium (BayGenomics)

Email: info@baygenomics.ucsf.edu

Sequence tag generated by 5' RACE of total RNA from gene trap ES

cell line. ES cell lines harboring insertion mutation of target

gene are available upon request from BayGenomics. Annotation

information available from

http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=

CELL_LINE&KEY=XE464

Class: Gene trap.

Location/Qualifiers

1..94

/organism="Mus musculus"

/mol_type="mRNA"

/strain="129 Ola"

/db_xref="taxon:10090"

/sex="Male"

FEATURES
source

/cell_type="Embryonic stem cell"
 /clone_lib="BayGenomics Gene Trap Library pGTILxf"
 /note="Vector: pGTILxf"

ORIGIN

Alignment Scores:
 Pred. No.: 0.00165 Length: 94
 Score: 168.00 Matches: 31
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.68% Indels: 0
 DB: 8 Gaps: 0

US-09-674-237B-3 (1-1213) x CC199586 (1-94)

QY 210 ProProAlaAglutirPalavalProGlnSerSerArgLeuLysTyrArgGlnLeuPhe 229
 Db 1 CTTCCAGCAGCAGATGGCTGTGCTCAGTCATCAAGCTGAATACAGGCAGTTATTC 60
 QY 230 AsnSerHisAspLysThrMetSerGlyHisLeu 240
 Db 61 AACAGCCACGACAAACTATGAGTGACACTTA 93

RESULT 2

AA589590 79 bp mRNA linear EST 16-SEP-1997
 LOCUS
 DEFINITION v149c09.s1 Stragatene mouse skin (#937313) Mus musculus cDNA clone
 IMAGE:975568 3', similar to SW:YFJ4 YEAST P43603 HYPOTHETICAL 40.4
 KD PROTEIN IN PES4-HIS2 INTERGENIC REGION. ; mRNA sequence.

ACCESSION AA589590
 VERSION AA589590.1 GI:2402970

KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 79)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE

The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:556296

Possible reversed clone: similarity on wrong strand

Seq primer: -28m13 rev1 ET from Amersham

High quality sequence stop: 1.

FEATURES

source

1. .79
 /location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:975568"
 /sex="females"
 /tissue_type="whole skin"
 /dev_stage="11 weeks old"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_lib="Stratagene mouse skin (#937313)"
 /note="Organ: skin; Vector: pBluescript SK-; Site: 1:
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
 Oligo dT. Whole skin from 11 week old C57BL/6 female mice.
 Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'

ORIGIN

Alignment Scores:
 Pred. No.: 3.16 Length: 79
 Score: 122.00 Matches: 24
 Percent Similarity: 96.15% Conservative: 1
 Best Local Similarity: 92.33% Mismatches: 1
 Query Match: 1.95% Indels: 0
 DB: 1 Gaps: 0

US-09-674-237B-3 (1-1213) x AA589590 (1-79)

QY 1002 AlaMetTyrThrTyrGluSerSerGluGlnGlyAspLeuThrPheGlnGlnGlyAspVal 1021
 Db 2 GCCATGTACACATACGAGAGTTCTGATCAAGAGATTAAAGTTTCAGCAGGGGATGTG 61

QY 1022 IleValValThrLysLys 1027

Db 62 ATTGTGTTACCAAGAA 79

RESULT 3

AV963382

LOCUS

DEFINITION AV963382 Nori Satoh unpublished cDNA library, egg Ciona
 intestinalis cDNA clone cig21112 5', mRNA sequence.

ACCESSION AV963382

VERSION AV963382.1 GI:19451681

KEYWORDS EST.

SOURCE

Ciona intestinalis
 Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Ciona.

1 (bases 1 to 99)

Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.

Expressed genes in Ciona intestinalis

Unpublished (2000)

Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoh@acidian.zool.kyoto-u.ac.jp.

FEATURES

source

1. .99
 /location/Qualifiers
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="cig21112"
 /tissue_type="whole animal"
 /dev_stage="egg"
 /clone_lib="Nori Satoh unpublished cDNA library, egg"

ORIGIN

Alignment Scores:

Pred. No.: 399 Length: 99
 Score: 95.50 Matches: 16
 Percent Similarity: 65.52% Conservative: 3
 Best Local Similarity: 55.17% Mismatches: 5
 Query Match: 1.52% Indels: 5
 DB: 2 Gaps: 1

US-09-674-237B-3 (1-1213) x AV963382 (1-99)

QY 1099 LysAsnProGlyGlyTyrTrpGluGluGlnAlaArgLysLysArgGlnIle 1118
 Db 5 AAGACAGAGGTGATGCTGGAGGGAATAACATCGAAAA-----ATT 49

QY 1119 GlyTrpPheProAlaAsnTyrVallys 1127

Db 50 GGTGGTTCCTCCACGCACTATGTGAAG 76

REFERENCE
1 (bases 1 to 98)
AUTHORS Schmid, K.J., Soerensen, T.R., Stracke, R., Torjek, O., Altmann, T.,
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

ORIGIN					
Alignment Scores:					
Pred. No.:	1.64e+03	Length:	97	/tissue_type="Stem"	
Score:	87.00	Matches:	17	/clone_lib="ST"	
Percent Similarity:	71.88%	Conservative:	6		
Best Local Similarity:	53.12%	Mismatches:	9		
Query Match:	1.39%	Indels:	0		
DB:	1	Gaps:	0		
US-09-674-237B-3 (1-1213) x AU250968 (1-97)					
QY	413	GlulysGlnArgGluLeuGluArgGlnArgGluGluGluArgLysGluIleGluArg	432		
Db	1	GA	60		
QY	433	ArgGluAlaLaLysArgGluLeuGluArgGlnArg	444		
Db	61	GA	96		
RESULT 8					
AW691901					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					
FEATURES					
source					
ORIGIN					
Alignment Scores:					
Pred. No.:	1.64e+03	Length:	97		
Score:	87.00	Matches:	17		
Percent Similarity:	71.88%	Conservative:	6		
Best Local Similarity:	53.12%	Mismatches:	9		
Query Match:	1.39%	Indels:	0		

DB: 2 Gaps:

US-09-674-237B-3 (1-1213) x AW691901 (1-97)

QY 413 GluLysGlnArgGluLeuGluArgGlnArgGluGluGluArgGlyGluGluGluArg 432
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 GA 60

QY 433 ArgGluAlaLysArgGluLeuGluArgGlnArg 444
::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 GA 96

RESULT 9
AW693216 97 bp mRNA linear EST 21-DEC-2000
LOCUS NF061F12ST1F1000 Developing stem Medicago truncatula cDNA clone
DEFINITION NF061F12ST 5', mRNA sequence.
ACCESSION AW693216
VERSION AW693216
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

REFERENCE 1 (bases 1 to 97)
AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A.,
Beil,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and
Dixon,R.A.

TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL Medicago truncatula stem library
COMMENT Unpublished (2000)
On Apr 14, 2000 this sequence version replaced gi:7567952.
Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 856 Std Error: 0.00
Plate: 061 row: F column: 12
Seq primer: TCACACAGGAACAGCTATGAC.

FEATURES
source
Location/Qualifiers
1..97
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF061F12ST"
/tissue_type="stem"
/dev_stage="Pooled developmental"
/clone_lib="Developing stem"
/note="Vector: Lambda Zap; Contains a mixture of
internodal stem segments"

ORIGIN
Alignment Scores:
Pred. No.: 1.64e+03 Length: 97
Score: 87.00 Matches: 17
Percent Similarity: 71.88% Conservative: 6
Best Local Similarity: 53.12% Mismatches: 9
Query Match: 1.39% Indels: 0
DB: 2 Gaps: 0

US-09-674-237B-3 (1-1213) x AW693216 (1-97)

QY 414 LysGlnArgGluLeuGluArgGlnArgGluGluGluArgGlyGluGluGluArg 433
::::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 2 AGA 61

QY 434 GluAlaAlaLysArgGluLeuGluArgGlnArg 445
::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

0
0
4
3
0
1

```

/organism="Capsicum annuum"
/mol_type="mRNA"
/cultivar="Nokkwang"
/db_xref="taxon:4072"
/tissue_type="red pepper fruit pericarp"
/clone_lib="B303"
/molnote="pluescript II SK(+)XR; Site.1: EcoRI; Site.2: XhoI; cDNA library was generated from red ripe fruit pericarp using lambda Zap II phage vector. In vivo excision was done with helper phage to generate subclone in pluescript II SK(+)XR vector."

```

Alignment Scores:

Alignment Scores:	2.06e+03	Length:	90
Pred. No.:	85.00	Matches:	17
Score:	72.41%	Conservative:	4
Percent Similarity:	72.41%	Mismatches:	8
Best Local Similarity:	58.62%	Indels:	0
Query Match:	1.36%	Gaps:	0
Da:	7		

UIS-09-674-237B-3 (1-1213) x CO912202 (1-90)

[illegible]

Search completed: April 29, 2005, 13:01:47
Job time : 8540 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2005, 03:08:25 ; Search time 446 Seconds

(without alignments)
4450.232 Million cell updates/sec

Title: US-09-674-237B-3

Perfect score: 6269

Sequence: 1 MAQFPFPGSLDVWATVE.....VGLFPSNYKLTMDMPSQ 1213

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delgap 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 1330268

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-O=/cgn2_1/USPTO spool/US09674237/rnat 27042005 182122 2902/app_query.fasta.1.1351
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=szmlm100.rni -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=100
-USER=US09674237 @CNC 1.1.93 @runat 27042005 182122 2902 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	88	1.4	96	3	US-09-281-481A-6
C 2	88	1.4	96	3	US-09-281-481A-7
C 3	65.5	1.0	87	4	US-09-513-999C-18169
C 4	65	1.0	69	4	US-09-573-080A-449
C 5	64	1.0	82	4	US-09-621-976-12952
C 6	64	1.0	96	3	US-09-461-697-223
C 7	63.5	1.0	96	3	US-08-484-322-28
C 8	60	1.0	58	3	US-09-281-481A-5
C 9	60	1.0	100	1	US-08-145-705A-4
C 10	58	0.9	100	1	US-08-145-705A-3
C 11	57	0.9	97	1	US-08-182-175A-54
C 12	57	0.9	97	1	US-08-474-633A-63

C 13	57	0.9	97	3	US-08-823-771-63	Sequence 63, Appl
C 14	57	0.9	97	5	PCT-US92-06412-54	Sequence 54, Appl
C 15	56.5	0.9	85	2	US-08-332-766A-44	Sequence 44, Appl
C 16	56	0.9	56	3	US-09-281-481A-3	Sequence 3, Appl
C 17	56	0.9	58	3	US-08-860-038-13	Sequence 13, Appl
C 18	56	0.9	58	3	US-08-860-038-14	Sequence 14, Appl
C 19	56	0.9	58	3	US-09-580-923-13	Sequence 13, Appl
C 20	56	0.9	58	3	US-09-580-923-14	Sequence 14, Appl
C 21	56	0.9	90	4	US-09-866-108A-15688	Sequence 15688, A
C 22	56	0.9	97	1	US-08-136-277-8	Sequence 8, Appl
C 23	56	0.9	97	2	US-08-479-403-8	Sequence 8, Appl
C 24	56	0.9	97	3	US-08-835-734-8	Sequence 8, Appl
C 25	55.5	0.9	78	4	US-09-579-894-4	Sequence 4, Appl
C 26	55	0.9	58	4	US-08-956-171B-2841	Sequence 2841, Ap
C 27	55	0.9	58	4	US-08-781-986A-2841	Sequence 13843, A
C 28	55	0.9	97	4	US-09-513-999C-13843	Sequence 3, Appl
C 29	55	0.9	100	1	US-08-145-705A-3	Sequence 31160, A
C 30	54.5	0.9	78	4	US-09-513-999C-31160	Sequence 31160, A
C 31	54	0.9	78	3	US-09-461-697-229	Sequence 229, App
C 32	54	0.9	91	1	US-08-222-177A-166	Sequence 166, App
C 33	54	0.9	92	1	US-08-222-177A-430	Sequence 430, App
C 34	54	0.9	97	1	US-08-182-175A-38	Sequence 38, Appl
C 35	54	0.9	97	1	US-08-182-175A-42	Sequence 42, Appl
C 36	54	0.9	97	1	US-08-182-175A-46	Sequence 46, Appl
C 37	54	0.9	97	1	US-08-474-633A-35	Sequence 35, Appl
C 38	54	0.9	97	1	US-08-474-633A-39	Sequence 39, Appl
C 39	54	0.9	97	1	US-08-474-633A-55	Sequence 55, Appl
C 40	54	0.9	97	3	US-08-823-771-35	Sequence 35, Appl
C 41	54	0.9	97	3	US-08-823-771-39	Sequence 39, Appl
C 42	54	0.9	97	3	US-08-823-771-55	Sequence 55, Appl
C 43	54	0.9	97	5	PCT-US92-06412-38	Sequence 38, Appl
C 44	54	0.9	97	5	PCT-US92-06412-42	Sequence 42, Appl
C 45	54	0.9	97	5	PCT-US92-06412-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1

US-09-281-481A-6/c
; Sequence 6, Application US/09281481A
; Patent No. 6383747
; GENERAL INFORMATION:
; APPLICANT: DAWKINS, Roger L. and ABRAHAM, Lawrence J.
; TITLE OF INVENTION: GENETIC ANALYSIS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/281,481A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/893,971
; FILING DATE: 16-JUL-1997
; APPLICATION NUMBER: US 232,229
; FILING DATE: 29-APR-1994
; APPLICATION NUMBER: PK9279 (AU)
; FILING DATE: 01-NOV-1991
; APPLICATION NUMBER: PCT/AU92/00583
; FILING DATE: 30-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLIO, FRANK S
; REFERENCE/DOCKET NUMBER: 9279


```

Query Match: 1.02% Indels: 0
DB: 4 Gaps: 0

US-09-674-237B-3 (1-1213) x US-09-621-976-12952 (1-82)

Qy 419 GluArgGluArgGluGluGluArgAGLYsGluileGluArgGluGluAalaIaIyAArg 438
Db 5 GAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 64
Qy 439 GluLeuGluArgGlnArg 444
Db 65 AAGAAAGAGAGAGAGAGAGG 82

RESULT 6
US-09-461-697-223
; Sequence 223, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: LO, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 223
; LENGTH: 96
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-223

Alignment Scores:
Pred. No.: 3, 86e+03 Length: 96
Score: 64.00 Matches: 11
Percent Similarity: 82.14% Conservative: 12
Best Local Similarity: 39.29% Mismatches: 5
Query Match: 1.02% Indels: 0
DB: 3 Gaps: 0

US-09-674-237B-3 (1-1213) x US-09-461-697-223 (1-96)

Qy 406 LeuGluLeuGluArgGlnLeuGluArgGluArgGluGluArgGluGluGlu 425
Db 1 ATGAAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Qy 426 ArgArgLYsGluileGluArgArg 433
Db 61 AAGAGAAGGAGAGATATAAAAGAGG 84

RESULT 7
US-08-484-322-28
; Sequence 28, Application US/08484322
; Patent No. 6110465
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R. H. AND
; APPLICANT: PURCELL, R. H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF HYPERVARIABLE
; TITLE OF INVENTION: REGION 1 OF THE ENVELOPE 2 GENE OF ISOLATES
; TITLE OF INVENTION: OF HEPATITIS C VIRUS AND THE USE OF
; TITLE OF INVENTION: REAGENTS DERIVED FROM THESE HYPERVARIABLE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:

```



```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
;
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Candida albicans
;
; US-08-145-705A-3
;
; Alignment Scores:
; Pred. No.: 1.16e+04
; Score: 58.00
; Percent Similarity: 53.85%
; Best Local Similarity: 41.03%
; Query Match: 0.93%
; DB: 1
;
; US-09-674-237B-3 (1-1213) x US-08-1
;
; QY 144 ProProLeuValSerVal
;
; DB 11 CCTCTCTGCGCTTCTTCTTGG
;
; QY 162 AlaProProValIleGlnPro
;
; DB 71 TGGCGCT-----CCT
;
; RESULT 11
;
; US-08-182-175A-54/c
; Sequence 54, Application US/08182
; Patent No. 5559223
;
; GENERAL INFORMATION:
; APPLICANT: Saverio Carl Falco
; APPLICANT: Sharon J. Keeler
; APPLICANT: Janet A. Rice
; TITLE OF INVENTION: Synthetic
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. du Pont de
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
;
; ZIP: 19898
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft word, 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/1
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/743,
; FILING DATE: 9 August 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Linda Axamethy Floyd
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: BB
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 992-4929
; TELEFAX: (302) 892-7949
; TELEX: 835420
;
; INFORMATION FOR SEQ ID NO: 54:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; STRAIN: E. coli
; CELL TYPE: DH5 alpha
; CLONE: 92-2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..88
; OTHER INFORMATION: /function= "synthetic storage protein"
; OTHER INFORMATION: /product= "protein"
; OTHER INFORMATION: /gene= "ssp"
; OTHER INFORMATION: /standard_name= "5.11.11.5"
; US-08-182-175A-54
Alignment Scores:
Pred. No.: 1.31e+04 Length: 97
Score: 57.00 Matches: 14
Percent Similarity: 47.50% Conservative: 5
Best Local Similarity: 35.00% Mismatches: 11
Query Match: 0.91% Indels: 10
DB: 1 Gaps: 2

US-09-674-237B-3 (1-1213) x US-08-182-175A-54 (1-97)
QY 125 ProLeuThraValaProValProMetGlySerIleProValValGlyMetSerPro 144
Db 92 CCTATCAGC-----CCTTCATCTTCTTCTCCATCCACTTCATCTTCTCT 48

QY 145 ProLeuValSerValProProAlaAlaValProProLeuAlaAsnGlyAlaProPro 164
Db 47 CCATCCACTTCATCTTCTCTCCATCG-----CCTTCATCTTCTCTCTCCA 3

RESULT 12
US-08-474-633A-63/c
; Sequence 63, Application US/08474633A
; Patent No. 5773691
; GENERAL INFORMATION:
; APPLICANT: E. I. DU PONT DE NEMOURS AND COMPANY
; TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR INCREASING THE LYSINE AND THREONINE CONTENT
; TITLE OF INVENTION: INCREASING THE LYSINE AND THREONINE CONTENT
; TITLE OF INVENTION: OF THE SEEDS OF PLANTS
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD VERSION 2.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,633A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: BARBARA C. SIEGELL
; REGISTRATION NUMBER: 30,684
; REFERENCE/DOCKET NUMBER: BB-1037-C
; TELECOMMUNICATION INFORMATION:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; STRAIN: E. coli
; CELL TYPE: DH5 alpha
; CLONE: 92-2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..88
; OTHER INFORMATION: /function= "synthetic storage protein"
; OTHER INFORMATION: /product= "protein"
; OTHER INFORMATION: /gene= "ssp"
; OTHER INFORMATION: /standard_name= "5.11.11.5"
; US-08-182-175A-54
Alignment Scores:
Pred. No.: 1.31e+04 Length: 97
Score: 57.00 Matches: 14
Percent Similarity: 47.50% Conservative: 5
Best Local Similarity: 35.00% Mismatches: 11
Query Match: 0.91% Indels: 10
DB: 1 Gaps: 2

US-09-674-237B-3 (1-1213) x US-08-474-633A-63 (1-97)
QY 125 ProLeuThraValaProValProMetGlySerIleProValValGlyMetSerPro 144
Db 92 CCTATCAGC-----CCTTCATCTTCTTCTCCATCCACTTCATCTTCTCT 48

QY 145 ProLeuValSerValProProAlaAlaValProProLeuAlaAsnGlyAlaProPro 164
Db 47 CCATCCACTTCATCTTCTCTCCATCG-----CCTTCATCTTCTCTCTCCA 3

RESULT 13
US-08-823-771-63/c
; Sequence 63, Application US/08823771
; Patent No. 6459019
; GENERAL INFORMATION:
; APPLICANT: E. I. DU PONT DE NEMOURS AND COMPANY
; TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR INCREASING THE LYSINE AND THREONINE CONTENT
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD VERSION 2.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,771
; FILING DATE: 24-Mar-1997
; CLASSIFICATION: <Unknown>
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/474,633
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BARBARA C. SIEGELL
; REGISTRATION NUMBER: 30,684
; REFERENCE/DOCKET NUMBER: BB-1037-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-4931
; TELEFAX: 302-773-0164
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; STRAIN: E. coli
; CELL TYPE: DH5 alpha
; IMMEDIATE SOURCE:
; CLONE: 92-2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..88
; OTHER INFORMATION: /function= "synthetic
; storage protein
; /product= "protein"
; /gene= "esp"
; /standard_name=
; "5.11.11.5"
; SEQUENCE DESCRIPTION: SEQ ID NO: 63:
US-08-823-771-63

Alignment Scores:
Pred. No.: 1.31e+04 Length: 97
Score: 57.00 Matches: 14
Percent Similarity: 47.50% Conservatives: 5
Best Local Similarity: 35.00% Mismatches: 11
Query Match: 0.91% Indels: 10
DB: 3 Gaps: 2

US-09-674-237B-3 (1-1213) x US-08-823-771-63 (1-97)

QY 125 ProLeuThraAlaValaProValProMetGlySerIleProValValGlyMetSerPro 144
Db 92 CCTATCACG-----CCTTCATCTCTCTCCATCCACCTTCATCTCTCTCT 48
QY 145 ProLeuValSerValProProAlaAlaValProProLeuAlaAsnGlyAlaProPro 164
Db 47 CCATCCACTTCATCTCTCTCCATCG-----CCTTCATCTCTCTCTCTCA 3

RESULT 14
PCT-US92-06412-54/c
; Sequence 54, Application PC/TUS9206412
; GENERAL INFORMATION:
; APPLICANT: Saverio Carl Falco
; APPLICANT: Sharon J. Keeler
; APPLICANT: Janet A. Rice
; TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing E
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. du Pont de Nemours and Company
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh System, 6.0

```

```

; SOFTWARE: Microsoft Word, 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06412
; FILING DATE: 19920807
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/743,006
; FILING DATE: 9 August 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Linda Axamethy Floyd
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: BB-1031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 992-4929
; TELEFAX: (302) 892-7949
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; STRAIN: E. coli
; CELL TYPE: DH5 alpha
; IMMEDIATE SOURCE:
; CLONE: 92-2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..88
; OTHER INFORMATION: /function= "synthetic storage protein
; OTHER INFORMATION: /product= "protein"
; OTHER INFORMATION: /gene= "esp"
; OTHER INFORMATION: /standard_name= "5.11.11.5"
; PCT-US92-06412-54

Alignment Scores:
Pred. No.: 1.31e+04 Length: 97
Score: 57.00 Matches: 14
Percent Similarity: 47.50% Conservatives: 5
Best Local Similarity: 35.00% Mismatches: 11
Query Match: 0.91% Indels: 10
DB: 5 Gaps: 2

US-09-674-237B-3 (1-1213) x PCT-US92-06412-54 (1-97)

QY 125 ProLeuThraAlaValaProValProMetGlySerIleProValValGlyMetSerPro 144
Db 92 CCTATCACG-----CCTTCATCTCTCTCCATCCACCTTCATCTCTCTCT 48
QY 145 ProLeuValSerValProProAlaAlaValProProLeuAlaAsnGlyAlaProPro 164
Db 47 CCATCCACTTCATCTCTCTCCATCG-----CCTTCATCTCTCTCTCTCA 3

RESULT 15
US-08-332-766A-44
; Sequence 44, Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
; APPLICANT: JEFFREYS, Alec J.
; APPLICANT: ARMOUR, John
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

Fri Apr 29 16:20:40 2005

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,766A
FILING DATE: 01-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9326052.9
FILING DATE: 21-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BIRD, Donald J.
REGISTRATION NUMBER: 25,323
REFERENCE/DOCKET NUMBER: 217211/M94/0434/CB
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-332-766A-44

Alignment Scores:
Pred. No.: 1.17e+04 Length: 85
Score: 56.50 Matches: 12
Percent Similarity: 80.77% Conservative: 9
Best Local Similarity: 46.15% Mismatches: 4
Query Match: 0.90% Indels: 1
DB: 2 Gaps: 1

US-09-674-237B-3 (1-1213) x US-08-332-766A-44 (1-85)

QY 413 GluLysGlnArgGluLeuGluArgGlnArgGlu--GluGluArgArgLysGluIleGlu 431
Db 2 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 61
QY 432 ArgArgGluAlaAlaLys 437
Db 62 AAAAGGAAGGAATGAAA 79

Search completed: April 29, 2005, 13:09:14
Job time : 449 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2005, 06:53:05 ; Search time 2004 Seconds
(without alignments)
3684.812 Million cell updates/sec

Title: US-09-674-237B-3

Perfect score: 6269

Sequence: 1 MAQPTFPFGSLDVMAITVE.....VGLFPFNYKLTMDPSSQ 1213

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5642217 seqs, 3043843248 residues

Total number of hits satisfying chosen parameters: 5511450

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO_spool/US09674237/runat_27042005_182123_2929/app_query.fasta.1.1351
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=szmlm100.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=100 -USER=US09674237@cgn 1 1 767 @runat_27042005_182123_2929 -NCPU=6
-ICPU=3 -NO_MMAP -LARGQUERY_NEG_SCORES=0 -WAIT_DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTU5_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	102	1.6	60	10	US-09-908-975-10766	Sequence 10766, A
2	89	1.4	99	18	US-10-021-323-14338	Sequence 14338, A
3	84	1.3	96	18	US-10-021-323-14266	Sequence 14266, A
4	79	1.3	88	18	US-10-021-323-808	Sequence 808, App
5	79	1.3	89	18	US-10-021-323-14335	Sequence 14335, A
6	79	1.3	89	18	US-10-021-323-15706	Sequence 15706, A
7	77	1.2	88	19	US-10-488-936-2	Sequence 2, Appli
8	73	1.2	81	18	US-10-021-323-14096	Sequence 14096, A
9	72	1.1	99	15	US-10-153-244-24	Sequence 24, Appl
10	70	1.1	73	18	US-10-021-323-14326	Sequence 14326, A
11	70	1.1	76	18	US-10-021-323-15848	Sequence 15848, A
12	67.5	1.1	96	16	US-10-061-201-11	Sequence 11, Appl
13	66	1.1	98	17	US-10-449-831A-113	Sequence 113, App
14	65	1.0	69	10	US-09-854-867-449	Sequence 449, App
15	65	1.0	69	19	US-10-786-970A-449	Sequence 449, App
16	64	1.0	73	9	US-09-263-959-724	Sequence 724, App
17	64	1.0	96	9	US-09-922-261-223	Sequence 223, App
18	63	1.0	60	10	US-09-908-975-12208	Sequence 12208, A
19	62	1.0	86	14	US-10-085-906-174	Sequence 174, App
20	61.5	1.0	60	10	US-09-908-975-8038	Sequence 8038, Ap
21	60	1.0	92	16	US-10-029-386-22483	Sequence 22483, A
22	59	0.9	77	9	US-09-925-302-437	Sequence 437, App
23	59	0.9	77	10	US-09-925-302-437	Sequence 437, App
24	58	0.9	59	18	US-10-021-323-824	Sequence 824, App
25	58	0.9	60	18	US-10-021-323-14327	Sequence 14327, A
26	58	0.9	74	17	US-10-355-716-114	Sequence 114, App
27	58	0.9	78	17	US-10-449-831A-114	Sequence 114, App
28	58	0.9	90	16	US-10-029-386-22401	Sequence 22401, A
29	58	0.9	98	16	US-10-029-386-22401	Sequence 22401, A
30	57	0.9	65	10	US-09-908-975-27220	Sequence 27220, A
31	57	0.9	65	10	US-09-908-975-27975	Sequence 27975, A
32	57	0.9	87	18	US-10-333-894A-6	Sequence 6, Appli
33	57	0.9	92	9	US-09-864-761-21177	Sequence 21177, A
34	57	0.9	97	14	US-10-023-056A-63	Sequence 63, Appl
35	57	0.9	97	18	US-10-804-678-63	Sequence 63, Appl
36	56.5	0.9	98	9	US-09-864-761-22804	Sequence 22804, A
37	56	0.9	51	18	US-10-021-323-14331	Sequence 14331, A
38	56	0.9	58	16	US-10-275-071-13	Sequence 13, Appl
39	56	0.9	58	16	US-10-275-071-14	Sequence 14, Appl
40	56	0.9	65	10	US-09-908-975-27864	Sequence 27864, A
41	56	0.9	90	9	US-09-866-108-15688	Sequence 15688, A
42	56	0.9	90	18	US-10-723-361-15688	Sequence 15688, A
43	55	0.9	58	8	US-08-781-988A-2841	Sequence 2841, Ap
44	55	0.9	58	17	US-10-329-624-2841	Sequence 2841, Ap
45	55	0.9	72	14	US-10-015-535-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-09-908-975-10766
; Sequence 10766, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT FILING DATE: 2001-07-20
; CURRENT APPLICATION NUMBER: US/09/908,975
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337

uArg 432

[illegible]

```

; TITLE OF INVENTION: higher molecular weight entities and uses thereof
; FILE REFERENCE: 2385978
; CURRENT APPLICATION NUMBER: US/10/449,831A
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: USSN 60/384878
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 237
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 113
; LENGTH: 98
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-449-831A-113

Alignment Scores:
Pred. No.: 1.23e+04 Length: 98
Score: 66.00 Matches: 16
Percent Similarity: 85.19% Conservative: 7
Best Local Similarity: 59.26% Mismatches: 2
Query Match: 1.05% Indels: 2
DB: 17 Gaps: 2

US-09-674-237B-3 (1-1213) x US-10-449-831A-113 (1-98)

QY 413 GlysGlnArgGluLeuGluArgGlnArgGluGluArgGluArgGluLeuGluArg 432
Db |||||:||||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
82 GAGAAGAGAGAGAGAA--GAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 29
QY 433 ArgGluAlaAlaLysArgGlu 439
Db |||||:|||||:|||||:|||||
28 AGAGAGAGAGAGAGAGAGAA 8

RESULT 14
US-09-854-867-449
; Sequence 449, Application US/09854867
; Publication No. US20030224356A1
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL H
; APPLICANT: ROGAN, PETER K
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/854,867
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 613
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 449
; LENGTH: 69
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: repetitive sequence
; OTHER INFORMATION: many eutherial genomes. Length of core repeating element
; OTHER INFORMATION: ble and is often polymorphic
US-09-854-867-449

Alignment Scores:
Pred. No.: 9.52e+03 Length: 69
Score: 65.00 Matches: 12
Percent Similarity: 78.26% Conservative: 6
Best Local Similarity: 52.17% Mismatches: 5
Query Match: 1.04% Indels: 0
DB: 10 Gaps: 0

US-09-674-237B-3 (1-1213) x US-09-854-867-449 (1-69)

QY 410 LysGlnLeuGlysGlnArgGluLeuGluArgGlnArgGluGluArgGluArgGlu 429
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 430 IleGluArg 432
Db |||||
61 AGAGAGAGAG 69

```

Search completed: April 29, 2005, 13:42:45
Job time : 2007 secs